Studies in Computational Intelligence 461

Mukul Sarkar
Albert Theuwissen

A Biologically Inspired CMOS Image Sensor

Springer

VISIT…

Studies in Computational Intelligence 461

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/7092

Mukul Sarkar and Albert Theuwissen

A Biologically Inspired CMOS Image Sensor

Springer

Authors
Dr. Mukul Sarkar
Indian Institute of Technology Delhi
Hauz Khas
New Delhi
India

Prof. dr. ir. Albert Theuwissen
Harvest Imaging
Bree
Belgium

ISSN 1860-949X
e-ISSN 1860-9503
ISBN 978-3-642-34900-3
e-ISBN 978-3-642-34901-0
DOI 10.1007/978-3-642-34901-0
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012952472

© Springer-Verlag Berlin Heidelberg 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my family

Contents

1 Introduction 1
1.1 Perception of Vision 1
1.2 Motivation 4
1.2.1 Wide Field of View 6
1.2.2 Motion Detection 6
1.2.3 High Sensitivity to Low Light Intensity 7
1.2.4 Polarization 7
1.3 Book Organization 8
References 10

2 Natural and Artificial Compound Eye 13
2.1 Natural Compound Eye 13
2.1.1 Apposition Eye 15
2.1.2 Superposition Eye 16
2.2 Properties of Compound Eye 18
2.3 Polarization Vision in Compound Eyes 26
2.4 Artificial Compound Eye 28
2.4.1 Design Consideration of an Artificial Compound Eye 32
2.4.2 Design of Artificial Compound Eye 36
2.5 Micro-optics Design 37
2.5.1 Image Mapping and Distortion Correction 43
2.6 Summary 46
References 47

3 Design of a CMOS Image Sensor 49
3.1 Introduction 49
3.2 Photodetector 51
3.3 Basic CMOS Image Sensor Architecture 57
3.3.1 Passive Pixel Sensor 57
3.3.2 Active Pixel Sensor 58
3.3.3 Digital Pixel Sensor 62

3.4 CMOS Image Sensor Performance . . . 64
3.4.1 Quantum Efficiency and Spectral Responsivity . . . 64
3.4.2 Dynamic Range and Signal-to-Noise Ration (SNR) . . . 66
3.4.3 Pixel Conversion Gain . . . 68
3.4.4 Dark Current . . . 68
3.4.5 Noise in CMOS Image Sensor . . . 70
3.5 CMOS Image Sensor Peripherals . . . 73
3.5.1 Addressing . . . 73
3.5.2 Column Processing Circuits . . . 74
3.5.3 Analog-to-Digital Conversion . . . 76
3.6 Designed Sensor Overview . . . 78
3.6.1 Pixel Architecture . . . 79
3.6.2 Pixel Operation . . . 81
3.6.3 Pixel Layout . . . 87
3.6.4 Row/Column Addressing Logic . . . 88
3.6.5 Analog Signal Chain . . . 91
3.6.6 Digital Signal Chain . . . 95
3.7 Image Sensor Characterization . . . 97
3.7.1 Test Setup . . . 98
3.7.2 Measurements . . . 99
3.8 Summary . . . 102
References . . . 103

4 Design of a CMOS Polarization Sensor . . . 105
4.1 Polarization Vision . . . 105
4.2 Polarization of Light - Basics . . . 106
4.2.1 Polarization of Light from an Unpolarized Beam . . . 108
4.2.2 Polarization of Light - Representation . . . 111
4.3 Polarization Cameras . . . 115
4.4 Wire-grid Polarizer . . . 118
4.4.1 Fabrication of Wire-grid Polarizer . . . 122
4.4.2 Transmittance Efficiency and Extinction Ratio of Wire-grid Polarizers . . . 124
4.5 Design of a Polarization Sensor . . . 137
4.6 Polarization Measurements . . . 139
4.6.1 Measurement Setup . . . 139
4.6.2 Analog Polarization Measurements . . . 142
4.7 Wavelength Selection Using Metal Grid . . . 145
4.8 Summary . . . 152
References . . . 154

5 Material Classification Using CMOS Polarization Sensor . . . 157
5.1 Introduction . . . 157
5.2 Polarization and Fresnel Coefficients . . . 159

5.2.1 Polarization Properties of a Reflected Light . . . 160
5.2.2 Fresnel Reflectance Model . . . 164
5.3 Material Classification Measurements . . . 166
5.3.1 Measurement Setup . . . 167
5.3.2 Polarization Transmittance . . . 167
5.3.3 Material Classification Using the Degree of Polarization . . . 170
5.3.4 Material Classification Using the Stokes Parameters . . . 172
5.3.5 Material Classification Using Polarization Fresnel Ratio . . . 174
5.4 Metal Classification . . . 177
5.4.1 Metal Classification Using PFR . . . 179
5.4.2 Metal Classification Using Degree of Polarization . . . 181
5.5 Summary . . . 182
References . . . 184

6 Navigation Using CMOS Polarization Sensor . . . 185
6.1 Introduction . . . 185
6.2 Celestial Compass Based on Skylight Polarization . . . 188
6.3 Navigation Using Polarized Light by Insects . . . 191
6.4 Navigation Using Polarized Light for Autonomous Agents . . . 193
6.5 Polarization Based Compass . . . 197
6.5.1 Measurement Setup . . . 197
6.5.2 Measurement Results . . . 197
6.6 Incoming Light Ray Direction Detection and Sun Position Detection . . . 202
6.6.1 Measurement Setup . . . 202
6.6.2 Measurement Results for Incoming Light Ray Directions . . . 203
6.7 Real-Time Implementation of Navigation Compass . . . 207
6.8 Summary . . . 211
References . . . 212

7 Motion Detection and Digital Polarization . . . 215
7.1 Motion Detection . . . 215
7.1.1 Motion Detection - Models . . . 218
7.2 Motion Detection - Differential Imaging . . . 220
7.3 Motion Detection - Optical Flow . . . 223
7.3.1 Motion in Vertical Direction - Collision Detection . . . 224
7.3.2 Motion in Horizontal Direction . . . 231
7.4 Illumination Invariant and High Dynamic Range Motion Detection . . . 233
7.5 Digital Polarization . . . 239
7.6 Summary . . . 242
References . . . 243

8 Future Works 247
8.1 Future Works 247
8.1.1 Wide Field of View Imaging System with Polarization Sensitivity 247
8.1.2 High Angular Resolution Imaging System with Polarization Sensitivity 249
References 252

Summary 253

Acknowledgments 257

Chapter 1
Introduction

Visual perception is the ability to detect light and interpret it. The early explanation of vision was provided by two major ancient Greek schools of thought. One believed in the "emission theory" championed by scholars like Euclid and Ptolemy, according to which vision occurs when light rays emanate from the eyes and intercepted by visual objects. The other school championed by scholars like Aristotle and Galen, believed in "intromission" where vision occurs by something entering the eyes representative of the object [1]. The Persian scholar Ibn al-Haytham ('Alhazen') is credited for refining the intromission theory into the modernly accepted theory of perception of vision [2]. In his most influential "Book of Optics", he defines vision to be due to the light from objects entering the eye [3], [4].

Perception is unavoidably selective; one cannot see all there is to see. Visual perception is not merely a translation of retinal stimuli. As John Berger and Jean Mohr [5] quote "A photograph quotes from appearances, but in quoting, simplifies them". The primary objective of this book is to simplify the perceived visual world (i.e. the photograph) for machine vision applications.

This chapter presents a short overview of various perceptions of vision prevalent in different optical vision systems. The motivation and challenges to design an optical vision system based on the compound eye of the insects as opposed to the more conventional approach of mimicking single aperture eyes is presented.

1.1 Perception of Vision

Objects become visible through many different phenomena of light example reflection, refraction. The way in which creatures see differs in regard to their shape, perception, color visualization, resolution and depth perception. Most animals see the world in fuzzy shades of gray while some animals can see in total darkness, or even see colors beyond the human visual spectrum. Hunting birds use binocular vision to spot prey from thousands of feet above. Dogs and cats are color blind but have better peripheral and night vision than humans and are more sensitive to movements. The sensitivity of owls is 50-1000 times greater to low light intensities

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 1–11.
DOI: 10.1007/ 978-3-642-34901-0_1 © Springer-Verlag Berlin Heidelberg 2013

than unaided human night vision. Snakes have thermal pits in addition to normal eyes extending their spectral sensitivity into the infrared. Horses and zebras have their eyes pointing sideways giving them outstanding peripheral vision thus warning them of the presence of predators.

Eyes can be simply defined as organs or visual systems for spatial vision. There are fundamentally two different ways in which high resolution spatial vision can be achieved: either by increasing the number of photoreceptors in the visual system or by multiplying the visual system in its entirety [6]. These two alternatives lead to simple or single chambered eyes and compound eyes, respectively. The human eye is a single chambered eye with a lens and works much like a camera as shown in figure 1.1(a). Animals with such eyes comprise less than 6% of the species in the animal kingdom. More than 77% of the known animal species are insects and crustaceans with compound eyes as shown in figure 1.1(b). Such compound or complex eyes are composed of numerous simple single aperture eyes.

The performance of any eye is principally affected by three structural and two environmental features [7]. Among the former are (a) the angular spacing of the receptors, which determines the spatial resolution; (b) the quality of the optical structures; (c) the diameter of the photoreceptors. Among the latter are (a) the amount of light available to the receptors, and (b) motion (self or image) [8].

The term “resolution” is used in a loose way to mean “ability to resolve fine details” [8]. Resolving power or visual acuity is used to describe the smallest single object that an eye can detect. The two fundamental limitations to the resolving power or resolution of eyes are the wave (diffraction) and the particle (photon noise) nature of light. The wide lens aperture in humans provides a very fine resolution due to the narrow airy disc of a point object caused by diffraction. An airy disc is a diffraction pattern resulting from a uniformly illuminated circular aperture which has a bright region in the center with concentric rings around. The resolution of the small individual lenses of the compound eye is diffraction limited to about 1°.

The other limitation of the resolution is the photon noise. Each photoreceptor can only detect a finite number of quanta, and any such count is associated with an uncertainty. Effectively, this limits the ability to detect differences in luminance (contrasts) across the image. For a given size, an eye can split up the captured light between many receptor cells (high spatial resolution), or it can use the light to activate fewer receptor cells more strongly (high sensitivity). The luminance generated by natural light sources varies over eight orders of magnitude between sunlight and starlight [7]. In 1941, Hecht et al. [10] discovered that human vision is limited in dim light by the small numbers of available photons. At the absolute human threshold, each receptor receives one photon on an average for every 40 minutes and this is the same for insects [10]. The uncertainty in the detection of the number of quanta (photons), increases in low light conditions, thus degrading the visual signal-to-noise ratio [11], [12]. As there needs to be a certain amount of light captured in order to be reliably detected, the eyes of nocturnal animals tend to be optimized for sensitivity rather than spatial resolution [8]. The visual sensitivity of the compound eye is enhanced by the summation of the spatial information in the neighboring visual channels [13], [14], [15].

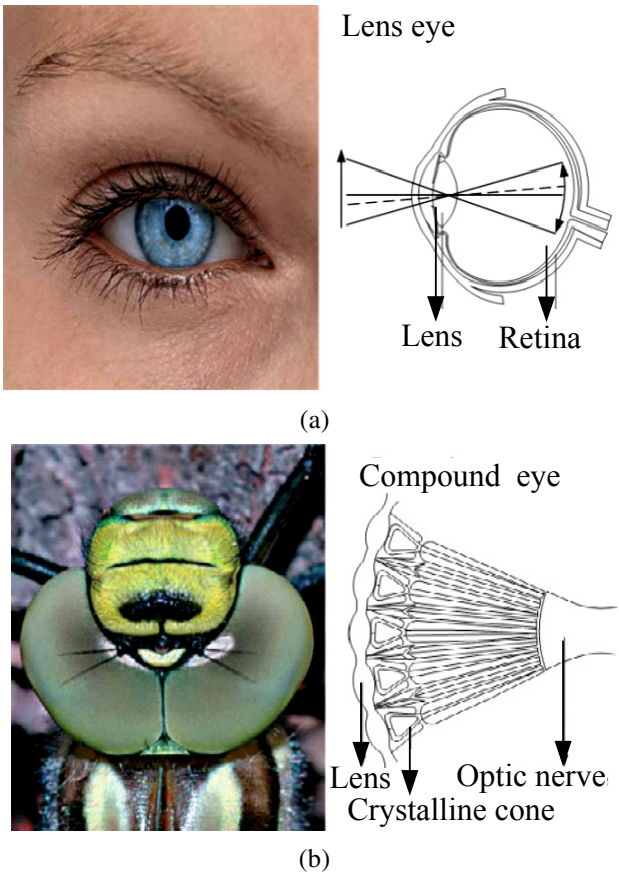


Fig. 1.1 (a) Single aperture eye, (b) Compound eye [9]

The visual acuity depends on both the spatial and the temporal resolution of the eye, as visual systems often detect and respond to objects that move relative to the animal itself [16]. As an object moves across the visual field, the individual receptor channels are progressively turned on and off due to a series of changes in the light intensity. This creates a *"flicker-effect"* known as the flicker frequency rate. The temporal resolving power for moving processes of a compound eye is considerably higher than for a single aperture eye, because of the high flicker frequency fusion rate. With their high speed of reaction, flying insects can resolve 250 images per second. In contrast the human eye sees only about 24 images per second. Eyes with high flicker fusion frequencies are able to recover responsiveness in shorter times than those with lower fusion frequency. Thus compound eyes are ideal for motion detection.

Motion blur is the smearing of a moving visual image due to the finite width of the impulse response of the photoreceptor. The combination of low spatial and high temporal resolutions makes the insects' visual system less sensitive to motion

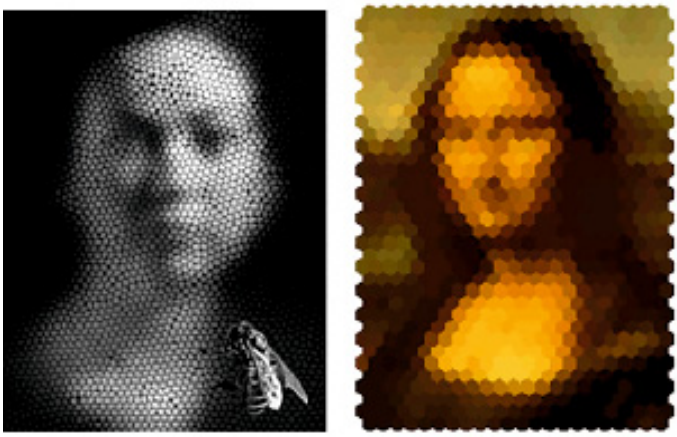


Fig. 1.2 An artistic impression of Insects (housefly's) vision [17]

blur than the human visual system. Theoretically, a fly's vision becomes blurred at an angular velocity of about 1.5°/7ms=200°/s, while in humans this occurs at 2'/25ms=1.4°/s. The wide field of view of the compound eye also aids in motion detection. The field of view of a compound eye is as close to 360° with respect to a two-dimensional plane, whereas the field of view of human eyes is around 160-208°.

Although an image of sorts forms in the eye, it is now accepted that for the majority of insects an image per se has no physiological significance. Figure 1.2 shows artistic impression of two images, as perceived by a housefly. The real function of the compound eye appears to be that of movement perception.

Besides being excellent in motion detection, compound eyes are also able to detect polarization. Light emanating from the sky and reflected light from a watery surface or shiny leaves is polarized i.e. it has greater vibration in some planes than in others. The microvillar organization of the insect's photoreceptors makes its visual systems inherently sensitive to the plane of polarization of light.

1.2 Motivation

The natural evolution of eyes to increase the spatial resolution by either increasing the number of photoreceptors in the visual system or by multiplying the visual system in its entirety has been of interest in the image sensor research field over the last 25 years. The increase of photoreceptors can be compared to the present day phenomenon of the "mega-pixel-race" and the design of autonomous multiple visual systems can be related loosely to the present day "neuromorphic image sensors". The neuromorphic approach implements specialized sensory processing function inspired by biological systems in analog electronic circuits. Neuromorphic vision sensors process images directly at the focal plane with circuits that try to emulate the first stages of visual processing in biological systems.

The perceptual world of animals is very different from that of humans as explained in section 1.1. Insects in general are exquisitely sensitive to image motion, which provides them with useful cues for avoiding obstacles and distance based landing. Thus insects' eyes are better suited for compact, robust and cheap vision systems.

The photoreceptor in a visual system is used to capture and convert the light photons to electric charge. The two most common semiconductor based photoreceptors used for artificial vision systems are CCD (charge coupled device) and CMOS (Complementary Metal Oxide Semiconductor) image sensors. CCDs produce a uniform output increasing the image quality, as all the pixels are only devoted to capture light. However CCDs usually have very limited output nodes, often only one for the pixel's charge to be converted into voltage to be sent off-chip. Also, CCDs have to be read out completely, and thus are inconvenient for region based readouts, affecting the output bandwidth of the sensor. They also require supporting circuits which increase the complexity and size of the camera design, thereby increasing the cost and power consumption. CMOS image sensors on the other hand use generic fabrication process, and in principle work very similarly to CCD sensors. In addition to being less expensive when compared to CCDs, they also allow for the integration of processing circuitry with the photoreceptor. CMOS image sensors also allow for random accessibility, enhancing the output bandwidth. The image quality of a CMOS image sensor is however not at par with that of a CCD as the in-pixel and in-column circuitry often produces a "fixed pattern noise" in the image.

Over the last decade, the resolution of the image sensor has been constantly increasing. Digital cameras with 12 Megapixels or more are now available in the market. To maintain the image quality with increasing resolution or decreasing pixel size, the hardware and firmware performance has to be increased, significantly affecting the complexity, size, cost and power consumption of the system.

Another area of research in vision systems has focused on neuromorphic or intelligent ways of processing the visual data. Until recently the artificial eye has been modeled based on the single aperture eye of the humans, wherein the output of the video-camera is scanned by a variety of spatial and temporal filters. This needs processing of huge volumes of data, defeating the purpose of having simple low power high sensitive vision sensors. Insects use minimal computations even for complicated functions like vision guided flying and thus are more suited for machine vision applications than single aperture eyes.

As we have discussed in the previous section, the compound eyes of insects offer low spatial resolution compared to single aperture eye but also offer other advantages for machine vision applications. These include:

- Wide field of view
- Better suited to detect moving object
- Higher sensitivity to light intensity
- Ability to detect polarization information

The focus of the book is to study each of the above mentioned advantages and to investigate new concepts and design techniques to implement the advantages of the

compound eye on conventional CMOS image sensors. Each of these advantages will be addressed in detail in this book.

1.2.1 Wide Field of View

A wide field of view is required in many applications like endoscopy, surveillance and automobile cameras. The most popular method to obtain this is image mosaicing, wherein multiple images are 'stitched' to obtain a larger image. This has a great redundancy in data, with each point being captured multiple times.

In comparison to the human eye which uses a spherical volume, compound eyes use only a spherical shell, permitting a large field of view but requiring less signal processing. Compound eyes have a field of view of near 360° obtained by the superposition of multiple simple lenses. Two such artificial insect facet eyes with this ability are the cluster eye and the artificial apposition compound eye *(APCO)* [18]. The cluster eye is inspired by the apposition eye and consists of three micro-lens array of different pitch. The optical axis of these micro-lenses are titled with respect to each other, and each micro-lens samples a small part of the object. The number of micro-lenses used determines the sensitivity and resolution. The *APCO* system is also based on the apposition eye where each channel samples one direction in space. The *APCO* system consists of a micro-lens array and a pinhole array with a slightly different pitch than the micro-lens array. The pinhole array defines the sensitivity and resolution. Both cluster eye and the *APCO* system has lower field of view *(FOV)* because the lenses are not tilted outwards, as is the case in insects.

The major technical drawback with wide angular field of view in artificial compound eyes is that the spherical optical world is mapped onto flat surfaces of the photo detector. The curved base of natural insect eyes offers several advantages including panoramic vision. Panoramic vision would mean far more information with which to monitor and control one's movement in the world. From a technological point of view, special fabrication techniques are being developed to have curved photodetectors, which would aid in the wide field of view vision [19].

1.2.2 Motion Detection

In conventional machine vision, motion is detected using a CMOS or CCD image sensor which samples the visual field at tens to thousands of times per second. This generates a huge volume of data to be processed for recognition and localization of the target. The precision of the system has a positive correlation with the computation payload. The higher the precision required, the higher is the payload.

The high flicker fusion frequency of the compound eye allows for much faster processing of the images than the single chambered eyes. The reaction time reported for honeybees to an object suddenly appearing in its visual field is 0.01s while that for humans is 0.05s. Additionally the division of the scene into multiple sub-images

allows parallel image processing and non sequential readout of the selected sub-images. This facilitates high speed object tracking while reducing image processing bottlenecks.

There are numerous architectures for motion detection based on the compound eye. Optical flow measurements have been explored by [20], [21], while frame differencing is used in [22]. Multi-resolution architecture has been proposed leading to variable spatial acuity imaging or region based motion detection [23], [24]. A CMOS image sensor with integrated circuits for basic decision making and image compression on the focal plane has also been presented [25]. The correlation motion detectors have been used to explain the direction selective motion using elementary motion detectors *(EMD)* [26]. One of the objectives of this book is to investigate on the algorithms to design a fast and simple motion detection algorithm with minimal complexity and processing.

1.2.3 High Sensitivity to Low Light Intensity

A real-world scene is composed of varying levels of brightness within it. The human eye has a very high dynamic range that enables it to detect subtle variation in the brightness and interpret scenes under different illumination conditions. However, while human vision is very poor in low light conditions some insects are known to have very good night vision. In dim light, to increase the visual signal-to-noise ratio, the visual system has to collect as much light as possible. The receptive field is optically enhanced in insects by summing the outputs of neighboring visual channels (spatial summation) or by increasing the sampling time also known as exposure period (temporal summation). The spatiotemporal summation can extend vision to light intensities more than 100,000 times dimmer than when relied on optics alone [27]. In apposition eyes, wider facets and wider photoreceptors can increase the sensitivity by one to two log units [28]. Receptor integration time may increase up to five fold in the dark. The enhancement of visual sensitivity by summation has negative effects on the spatial and temporal resolution.

As vision based safety measures for automotives eventually become mandatory in Europe starting 2013 [29], imaging innovations in low light vision systems will have to be in place. The current area of CMOS image sensors in the automotive industry focuses on high dynamic range and near infra-red type of image sensors. The dynamic range of an artificial optical system can be enhanced either by changes in the exposure period or with an adaptive photodetector [30]. These sensors coupled with night vision ability and more signal processing for intelligent decisions within the image sensor would be a value addition in assisted driving.

1.2.4 Polarization

The three basic characteristics of light are intensity, color and polarization. Polarization provides a more general description of light than either the intensity or the color alone, and can therefore provide richer sets of descriptive physical constraints for

the interpretation of the imaged scene. Further a polarization sensitive image sensor along with the ability to measure intensity and color will make it possible to sense the complete set of electromagnetic parameters of light incident on the camera.

The presence of linearly polarized light (the most common type of polarization in nature) in the optical environment can be qualitatively demonstrated by the use of a linear polarizer. The latest polarization image sensors utilize optical imaging systems that are external to the detectors [31]. These polarization filters are either single or multi-axis arrays which measure the polarization information in real time. The compactness of design and the speed at which polarized images are generated can be enhanced greatly by incorporating an array of microscopic polarization filtering optics directly onto a photosensitive chip. Micro-polarizers can be made either from organic materials [32] or using metallic wire grid available with the standard CMOS technology [33].

Light rays get reflected when they strike a reflecting surface, and the reflected light is polarized depending on the incident light and the nature of the reflecting surface. This polarized component can be used to detect the nature of the surface, for example to discriminate between a metal and a dielectric surface [34].

The polarized light in the optical environment is also known to be used by insects such as desert ants for navigation. The ability of the desert ant to navigate effortlessly in complex environments has been a subject of research in robotics [35]. The current vision sensors for navigation use a generalized algorithm of capturing two-dimensional images using a standard CMOS or a CCD camera and then processing those captured images for vector calculations to determine the motion and direction vector. This involves intensive processing along with high power consumption, as most of the post processing of the images is done in the digital domain on a FPGA or a microcontroller. Some of the cheap computational strategies that insects use to navigate have been already modeled in the past [36], [37].

In conclusion, the focus of this book can be summarized as

- Design of an image sensor with an intelligent or a smart pixel, to extend the functionality of the active pixel. The presence of a large number of active elements will limit the resolution. However as a large number of images can be processed at the pixel level, the overall output bandwidth of the image sensor will be improved. Furthermore pixel level processing will also reduce the amount of digital computations helping to design low power fast motion detectors.
- Design of a polarization detection sensor using a metallic wire grid using the metal layers available with a standard CMOS technology. By spatially orienting the micro-polarizer in varying directions polarization information from the optical world can be obtained. Further it is desired to show the real time detection and analysis of polarization information in machine vision applications like material classification and navigation.

1.3 Book Organization

This book consists of seven chapters. Chapter two begins with an overview of the compound eye found in the insects. The different types of the natural compound

eye like the apposition eyes and superposition eye are described. The properties of the compound eye like the visual acuity, interommatidial angle, angular sensitivity function, acceptance angle, filed of view, colour and spectral sensitivity are detailed. The polarization vision in the compound eyes of the insects is elaborated. The design of artificial compound eye is discussed along with available designs in literature. The design consideration for an artificial compound eye are presented, following which a description of the artificial compound eye designed in collaboration with the Vrije Universiteit Brussel *(VUB)* is presented.

Chapter three being with an introduction to the Charge-Coupled-Devices (CCD) and Complementary-Metal-Oxide-Semiconductor (CMOS) image sensors. The basic operation of the CMOS image sensor is further elaborated. The process of photodetection, and the available standard pixels operation are described in detail. The performance metrics of the CMOS image sensors are also elaborated. Further, the design and operation of the designed image sensor based on the compound eye of the insects is described. The performance characterization of the designed image sensor are also described.

Chapter four begins with an overview of the basic theory behind the polarization of light and its states. The absorption of the incident electromagnetic waves using specifically aligned metallic grid layers and its effect on the transmission intensity and wavelength are discussed using theoretical simulations. The micro-polarizer realized in the standard CMOS process and placed on top of the photodiode with varying transmission axis is described and characterized. The polarization information detected using the wire grid micro-polarizer can be used for many applications. Two such applications, material classification and navigation, are presented in chapter four and five respectively.

Chapter five begins with the overview of the Fresnel reflections of a polarized light ray after specular reflection. The metallic wire grid micro-polarizer in the designed image sensor allows computing parallel and perpendicular Fresnel reflection coefficients along with the maximum and the minimum transmitted irradiance after reflection from the material surface. Using these measured parameters, various methods described in the literature are used to test the ability of the polarized reflected light to distinguish between a metallic and a dielectric surface. The measured coefficients are further shown to allow for differentiating between various metallic surfaces based on conductivity.

Chapter six begins with the introduction to the natural navigation patterns as exhibited by insects like *Cataglyphis fortis*. Navigation using the polarized state of the incoming light ray is proposed. The azimuthal and the elliptical position of the incoming light rays can be detected using the polarized state of the incoming light ray. This can further be used to detect the position of the sun, as naturally available skylight is always polarized and the degree of polarization depends on the position of the sun.

Chapter seven presents the ability of the image sensor to detect motion based on optical flow and differential imaging. The one-dimensional binary optical flow is used to detect collision (vertical motion) and also horizontal motion. The dynamic range of a sensor plays an important role in detecting motion. Partial charge transfer

mechanism is shown to increase the dynamic range and also shows invariance to the changes in the background illumination. The one-dimensional optical flow varies with the polarization angle and can be used to present polarization in digital form.

References

[1] http://en.wikipedia.org/wiki/Visual_perception
[2] Sabra, A.: The Optics of Ibn al-Haytham. Warburg Institute, London (1989) ISBN: 0854810722
[3] Sabra, A.: Ibn al-haytham, brief life of an arab mathematician: died circa 1040. Harvard Magazine, 54–55 (2003)
[4] Sabra, A.: Ibn al-haytham's criticisms of ptolemy's optics. Journal of the History of Philolosophy 4, 145–149 (1966)
[5] Berger, J., Mohr, J.: Another Way of Telling. Pantheon, New York (1982)
[6] Nilsson, D.: The evolution of eyes and visually guided behaviour. Philosophical Transactions of the Royal Society London B: Biological Sciences 364(1531), 2833–2847 (2009)
[7] Land, M., Nilsson, D.: Animal eyes. Oxford University Press, Oxford (2002) ISBN: 0198575645
[8] Land, M.: Visual acuity in insects. Annual Review of Entomology 42(1), 147–177 (1977)
[9] Innovation. Carl Zeiss Magazine (17), 16–17 (2006)
[10] Pirenne, M.: Vision and the Eye. Chapman and Hall, London (1967)
[11] de Vries, H.: The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. Physica 10, 553–564 (1943)
[12] Rose, A.: The relative sensitivities of television pickup tubes, photographic film, and the human eye. Proceedings of the Institute of Radio Engineering 30(6), 293–300 (1942)
[13] Srinivasan, M., Laughlin, S., Dubs, A.: Predictive coding: A fresh view of inhibition in the retina. Philosophical Transactions of the Royal Society London B: Biological Sciences 216(1205), 427–459 (1982)
[14] Tsukamoto, Y., Smith, R., Sterling, P.: Collective coding of correlated cone signals in the retinal ganglion cell. Proceedings of the National Academy of Sciences America 87, 1860–1864 (1990)
[15] van Hateren, J.: A theory of maximizing sensory information. Biological Cybernetics 68(1), 23–29 (1992)
[16] Srinivasan, M., Bernard, G.: The effect of motion on visual acuity of the compound eye: a theoretical analysis. Vision Research 15, 515–525 (1975)
[17] http://www.admin.cam.ac.uk/news/press/dpp/2000061501
[18] Duparre, J., Eisener, M., Weible, K.: Miniaturized imaging systems. Microelectronic Engineering 67-68(1), 461–472 (2003)
[19] Ko, H.C., Stoykovic, M., Song, J., Malyarchuk, V., Choi, W., Yu, C., Gegges III, J.B., Xiao, J., Wang, S., Huang, Y., Rogers, J.: A hemispherical electronic eye camera based on compressible silicon optoelectronics. Nature 454, 748–753 (2008)
[20] Moini, A., Bouzerdoum, A., Eshraghian, K., Yakovleff, A., Nguyen, X., Blanksby, A., Beare, R., Abbott, D., Bogner, R.: An insect vision-based motion detection chip. Journal of Solid-State Circuits 32(2), 279–284 (1997)
[21] Mehta, S., Cummings, R.: Normal optical flow measurement on a CMOS APS imager. In: Proceedings of International Symposium on Circuits and Systems, vol. 4, pp. 848–851 (2004)

[22] Milirud, V., Fleshel, L., Zhang, W., Jullien, G., Pecht, O.: A wide dynamic range CMOS active pixel sensor with frame difference. In: Proceedings of International Symposium on Circuits and Systems, vol. 1, pp. 588–591 (2005)
[23] Kemeny, S., Panicacci, R., Pain, B., Matthies, L., Fossum, E.: Multiresolution image sensor. IEEE Transactions on Circuits and Systems for Video Technology 7(4), 575–583 (1997)
[24] Saffih, F., Hornsey, R.: Multiresolution CMOS image sensor. Technical Digest of SPIE Opto, Canada, pp. 425–428 (2002)
[25] Mallik, U., Clapp, M., Choi, E., Cauwenberghs, G., Cummings, R.: Temporal change threshold detection imager. In: IEEE International Solid-State Circuits Conference, Digest of Technical Papers, vol. 1, pp. 362–363 (2005)
[26] Borst, A.: Correlation versus gradient type motion detectors: the pros and cons. Philosophical Transactions of the Royal Society London B: Biological Sciences 362(1479), 369–374 (2007)
[27] Warrant, E.: Seeing better at night:life style, eye design and the optimum strategy of spatial and temporal summation. Vision Research 39(9), 1611–1630 (1999)
[28] Williams, D.: Changes of photoreceptor performance associated with the daily turnover of photoreceptor membrane in locusts. Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology 150, 509–519 (1983)
[29] Robin, L., Baron, J.: CMOS image sensors technologies and market - 2010 report. Tech. rep., Market research reports, Yole développement (2010)
[30] Darmont, A.: Methods to extend the dynamic range of snapshot active pixels sensors. In: Proceedings of SPIE, International Society of Optical Engineering, vol. 6816, pp. 681603.1–681603.11 (2008)
[31] Andreou, A., Kalayjian, Z.: Polarization imaging: principles and integrated polarimeters. IEEE Sensors Journal 2(6), 566–576 (2002)
[32] Zhao, X., Boussaid, F., Bermak, A., Chigrinov, V.: Thin photo-patterned micropolarizer array for CMOS image sensors. IEEE Photonics Technology Letters 21(12), 805–807 (2009)
[33] Tokuda, T., Yamada, H., Sasagawa, K., Ohta, J.: Polarization analyzing CMOS image sensor with monolithically embedded polarizer for microchemistry systems. IEEE Transactions on Biomedical Circuits and Systems 3(5), 259–266 (2009)
[34] Wolff, L.: Polarization based material classification from specular reflection. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(11), 1059–1071 (1990)
[35] Wehner, R., Michel, B., Antonsen, P.: Visual navigation in insects: coupling of egocentric and geocentric information. Journal of Experimental Biology 199, 129–140 (1996)
[36] Lambrinos, D., Maris, M., Kobayashi, H., Labhart, T., Pfeifer, R., Wehner, R.: An autonomous agent navigating with a polarized light compass. Adaptive Behaviour 6(1), 131–161 (1997)
[37] Usher, K., Ridley, P., Corke, P.: A camera as a polarized light compass: preliminary experiments. In: Proceedings of Australian Conference on Robotics and Automation, pp. 116–120 (2001)

Chapter 2
Natural and Artificial Compound Eye

The compound eye is one of the most compact vision systems found in nature. The eye is a compound of individual lenses, each one with their own photoreceptor arrays. This visual system allows an insect to fly with a limited intelligence and brain processing power. In comparison to single-aperture eyes, compound eyes have much wider field of view *(FOV)*, better capability to detect moving objects, higher sensitivity to light intensity, but much lower spatial resolution.

Section 2.1 of this chapter discusses the nature and types of natural compound eye of insects. The apposition and superposition compound eyes are discussed. The properties of the compound eye like the visual acuity, interommatidial angle, angular sensitivity function, acceptance angle, filed of view, colour and spectral sensitivity are discussed in detail in section 2.2. In section 2.3 the ability of the compound eyes to detect polarized light and its use for navigation and identification of food/communication is discussed. Human eyes are polarization blind and only detect it as a glare. The lens designers and manufactures have struggled long to attain a balance between the image quality and the wide angle lenses. On the other hand the insects with compound eyes are able to achieve extremely wide field of view. Section 2.4 describes the design flow of an optical lens array inspired by the multiple simple lens structure of the compound eye. Section 2.5 describes the designed micro-optical lens array briefly. The mapping of the spherical world into the planar structure introduces non-uniformities. The radial and photometric non-uniformities are elaborated and methods are proposed to remove these non-uniformities.

2.1 Natural Compound Eye

Eyes can be neatly divided into two major groups, simple and compound eyes. The purpose of the eye is to provide information to the nervous system about the light intensity at different angles. The raw image is basically the same for both types of eyes. The quality of the raw image is determined by basically two factors; the optical image quality and the density of the array of photoreceptors cells in the

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 13–48.
DOI: 10.1007/ 978-3-642-34901-0_2 © Springer-Verlag Berlin Heidelberg 2013

eye that forms the image [1]. The insects eyes are formed by bending in a convex direction (outward) as compared to the cup-like eye of the humans, which evolved through bending of the photoreceptors concavely (inward).

The compound eye of an insect is a faceted structure consisting of individual, radially-arranged optical channel units called ommatidia (singular ommatidium) as shown in figure 2.1. Each ommatidium appears on the surface as a single polygon or dome and is called a facet. The facets can be of many shapes like hexagonal, squarish, circular or hemispherical. The packing density is the highest for hexagonal facets on the surface of the eye. Each ommatidium captures light from a certain cone of angle of incidence contributing a spot to the visual image, which consists of the total inputs from all the ommatidia in the eye resulting in a mosaic like picture of the world. Thus each individual lens allows the insects to see a part of the overall image. At the micro level, compound eyes individual viewing channels on curved lenses have minimal aberrations, or focusing errors that cause blurring.

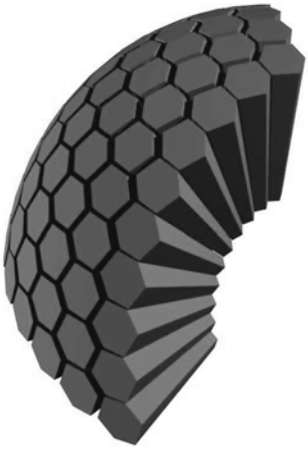


Fig. 2.1 Schematic of an compound eye

Each ommatidium as shown in figure 2.2 consists of a lens that forms an image onto the tip of the light sensitive visual cells *rhabdom*, a transparent crystalline cone to guide the light and absorptive pigments between the ommatidium [2]. Each ommatidium also contains seven or eight sensory cells, sometimes referred as retinula cells surrounding the rhabdom, where the photon absorbing, visual pigments are arranged within microvilli. The light sensitive parts of the visual cells are microvillis, which are an array of tubelike membranes containing the visual pigment rhodopsin. The visual pigments *(rhabdomeres)* are highly refractive materials, and are ideal for collecting the available light with minimal losses as the available light keeps getting totally internally reflected. Thus the rhabdom are ideal structure to receive and conduct light without any loss. These rhabdomeres can either be fused together as a rod-shaped rhabdom, or open (separated) throughout their length. This visual information is transduced into electrical signals by the

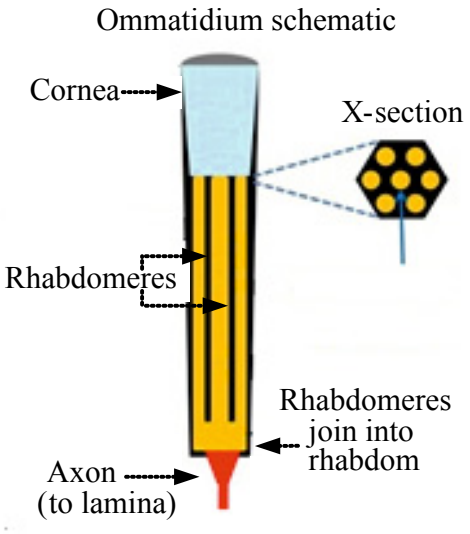


Fig. 2.2 Schematic of an ommatidium [3]

photoreceptor cells at the level of the retina. This transduced signal from the retina is processed by three layers of the visual ganglia: the lamina, the medulla, and the lobula complex.

Compound eyes can broadly be classified into two types: apposition and superposition types of compound eyes [4].

2.1.1 Apposition Eye

Apposition eyes are common in diurnal (day-active) insects and are most widely spread. Apposition eyes have a single positive lens, which forms an image by sampling a relatively large field of view of the scene. The image produced is not resolved in the retina rather by the visual pigment, rhabdom. In an apposition eye, each optical channel has its own receptor. The rhabdomeres and cones touch each other and there exists an absorptive screening pigment between the ommatidia. This reduces the amount of light that can reach a rhabdom from cones other than the one above it, thus keeping the angle of acceptance of light of each ommatidium low. In such eyes each optic unit represents only one sample point. The rhabdom measures the light intensity in the center of the image. The rhadbom has a very narrow field of view and resembles single rod in a vertebrate eye [1]. Apposition eyes are generally found in day-active insects such as ants, wasps, dragonflies, bees and cockroaches.

The apposition eyes consists of very small facet, about 10μm in diameter. Due to this small aperture, the apposition eyes work best at bright light intensities. The final image in apposition eye is made of discreet points, each point contributed by a

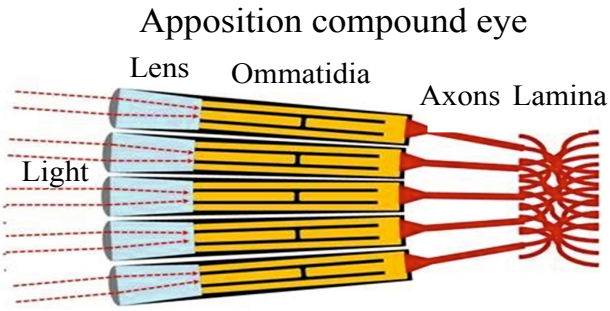


Fig. 2.3 Apposition eye [3]

single ommatidium, which appears as a mosaic of points. The information from the ommatidia are integrated in the brain to obtain a final image.

The apposition eyes have the disadvantage of having a small lens size. The small size of the lens diffracts the light larger and only little light reaches the rhabdoms. To improve the light sensitivity, in some flies, the rhabdom is split into multiple light guides. The increased number of light guides helps in capturing more the available light thus increasing the light sensitivity. The signals from the individual light guides are summed neurally to produce a stronger signal for the brain. This type of eye is often referred to as apposition-neural superposition eye.

2.1.2 Superposition Eye

In a superposition compound eye the light from different adjacent channels is guided to a common receptor. This greatly improves the probability of a photon capture and thus the sensitivity. The gain in sensitivity is approximately equivalent to the number of superimposed facets. The superposition eyes are typically more sensitive to light, and this type is found mainly in nocturnal animals that live under low light conditions, such as moths and crustaceans living in dim mid-water regions of the ocean. The optical superposition eyes are much superior compared to the apposition eyes. The rhabdoms of the superposition eyes are usually wider and shorter compared to the apposition eyes.

Superposition compound eyes can be subdivided into three types: refractive, reflective and neural as shown in figure 2.3. Refractive superposition eyes have an open space between the lens array and the rhabdom (array of photoreceptors). In refractive superposition eyes, a large lens cylinder exists which helps in the formation of an upright image as compared to the single positive lens in apposition eye. In insects eyes usually the lens in the cornea is powerful, which assist the crystalline cone. However, in refractive superposition the lens cylinder in the crystalline cone is much superior, in fact the shape of the lens cylinder is a characteristic of the refractive superposition eyes.

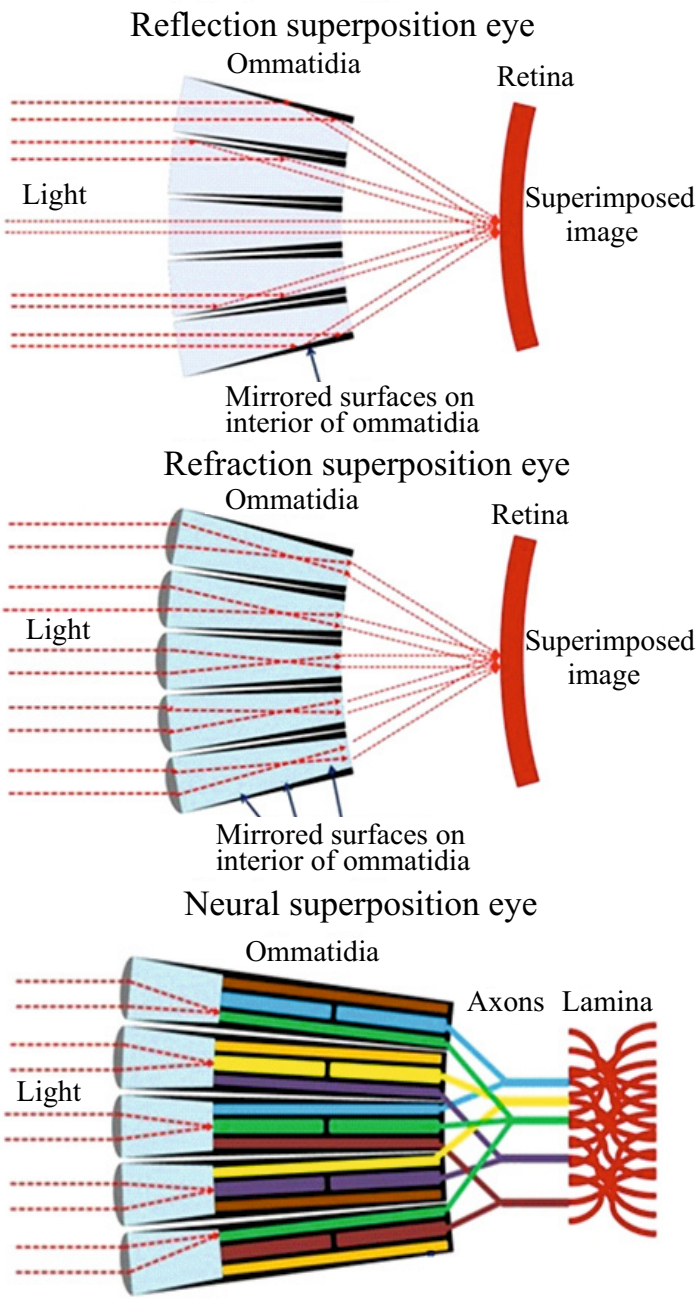


Fig. 2.4 Superposition eyes: reflection, refraction and neural type [3]

In reflective superposition eyes, the lenses are actually mirrors that reflect the light from a certain direction to a location on the rhabdom. The walls of the crystalline cone usually have a high refractive index and thus acts as mirrors. Further, these crystalline cones are sometimes lines with a multilayered reflector. The reflective superposition eyes don't need lenses, and the images are formed only using reflections in the mirrors. Instead of circular cross-section of the crystalline cone, the reflective superposition eyes have square cross-section of the crystalline cone, which limits the spreading of light rays in all directions.

In a neural superposition eye neural circuits puts together the signals from the individual small eye facets. The optical and neural organization of the eyes results in increased light sensitivity while lowering the spatial resolution due to neural pooling of information. Here each ommatidium receives light only through its own facet, thus the light sensitive structures, the rhabdomere, are optically isolated from each other. The retinula cells axons from the rhabdomere of adjacent ommatidia, which all have the same field of view converge together in the lamina. Thus each lamina effectively looks at only one point in space. This increases the sensitivity due to combined visual information from multiple ommatidia, however the spatial resolution is marginally decreased.

2.2 Properties of Compound Eye

The arrangement of the micro-lens layer and the detector layer in a natural apposition compound eye is shown in figure 2.5. The ommatidia are arranged on a curved surface with radius R. The micro-lens with diameter D and focal length f focuses light only from a small solid angle $\Delta\varphi$. The $\Delta\varphi$ is the acceptance angle of the individual optical channel. The receptor diameter is d. $\Delta\phi$ is the angular distance between the optical channel's axis also known as interommatidial angle. The ommatidia arranged on a spherical shell allows the compound eyes to have a very large field of view *(FOV)*.

Interommatidial Angle and Sensitivity

Visual acuity is used to describe the smallest single object that an eye can detect [5]. It is also defined as the partial resolution that the eye can see. Visual acuity is determined by the maximum possible resolution an eye can resolve. The resolving power of an eye is determined by the density of the retinal cells or ommatidia which constitutes the total number of image points in the final image. The fundamental determinant of acuity is the angle between the two detectors in an array. The visual acuity can be measured using interommatidial angle of the compound eye. The visual environment of the insect is sampled with the interommatidial angle, which is given by equation (2.1)

$$\Delta\phi = \frac{D}{R} \tag{2.1}$$

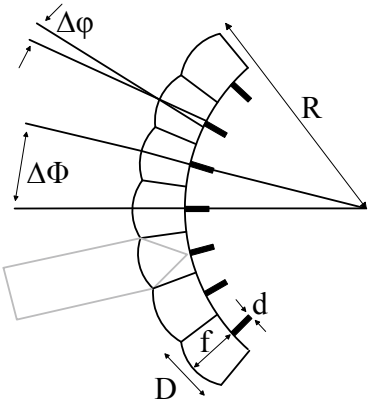


Fig. 2.5 Structure of a natural compound eye

where *D* is diameter of the micro-lens and *R* is radius of the eye. The photoreceptor responds to visual stimuli only if the object is located on the optical axis of an ommatidium. The sampling or Nyquist angular frequency with which an apposition eye can resolve line pattern of alternating bright and dark stripes is given by

$$v_s = \frac{1}{2\Delta\phi} \tag{2.2}$$

where v_s is the maximum possible visual acuity and $\Delta\phi$ is the interommatidial angle. The equation (2.2) also describes the maximum possible visual acuity of a compound eye. The visual acuity of the insects compound eye is much less (about one hundred times) compared to the human eye. The visual acuity depends both on the spatial resolution and temporal resolution of the eye, as the visual targets are rarely stationary [6]. Visual acuity is influenced by the optical limits and the imperfections in the lens system. The light passing through the lens will be diffracted which will cause blurring in the final image. The resolution of the insect eyes are severely limited by diffraction. The compound eye can be made large to achieve the same visual acuity as the human eye. Dragon flies among the insects have the largest compound eyes, several millimeters in diameter, for very sharp vision required to catch flying insects.

The visual acuity also depends on the density of the visual pigments present in the eye. In human eyes, the pitch of the sensor cells, rods and cones is only 2μm, which allows high packing density of the sensor cells. This allows human eye to have a higher resolution. The insects eye don't benefit from lenses in micrometer range due to diffraction limitation of the lenses. The smaller lens prevents light from being focussed into a point smaller than 5μm in diameter, thus small lenses don't help increasing the resolution. An eye can only resolve a finite amount of image details, because the wave and particle nature of the light will impose fundamental

limits on the information that an eye can provide. The wave property of light causes diffraction, which causes image blurring that is inversely proportional to the diameter of the opening through which light enters the eye. The minimum resolvable angle is given by

$$\theta = \frac{\lambda}{D} \tag{2.3}$$

where θ is the minimum resolvable angle (in radians); D is the diameter of the opening and λ is the wavelength of light. For a 3mm eye the minimum resolvable angle for a 0.5μm wavelength light is 0.0095°, while for a 50μm facets of the compound eye, it is 0.57°. The visual capacity increases with the eye size, however a larger eye is expensive in terms of its weight, volume, processing power and energy. The diffraction makes the compound eye with many small lenses inherently inferior to a single-lens eye. Generally the diurnal animals are diffraction limited, while nocturnal or deep sea animals are photon noise limited [7].

For a diffraction limited eye, the maximum visual acuity is given by

$$v_{s,l} = \frac{D}{\lambda} \tag{2.4}$$

where $v_{s,l}$ is the diffraction limited maximum visual acuity, D is the diameter of the ommatidium and λ is the wavelength of light. The wider the lens aperture, D, the narrower the airy disc, and thus finer the resolution in the image.

Using the equations (2.2) and (2.4), the radius of the curvature of the eye for maximum visual acuity is given by

$$R = \frac{\lambda}{2\Delta\varphi^2} \tag{2.5}$$

From equation (2.5), is it observed that the radius of the eye is proportional to the square of the required resolution. Thus for a ommatidium of diameter 20μm and interommatidial angle of 0.01°, the maximum achievable visual acuity for a 500nm light is

$$\Delta\phi = \frac{0.01}{\pi} = 1.745x10^{-4} \tag{2.6}$$

$$R = \frac{\lambda}{2\Delta\varphi^2} = \frac{500x10^{-9}}{2x(1.745x10^{-4})^2} = 8.2m \tag{2.7}$$

To have the spatial resolution close to human eye, the insect would need a eye of radius 8.2m, which is quite large.

Contrast is related to visual acuity in the sense of spatial resolution. It is the ability to distinguish similar shades of the same colour, and is important in defining the edges of an object. The contrast is dependent on the intensity of light, in dim light the contrast is less. Thus to increase the contrast, in dim light environment more light needs to be collected. There are two ways to increase the amount of light collected

- a bigger lens aperture allows more light to pass through it to reach the light sensor,
- the time for which the light is collected by the light sensor is increased

The animal in their daily life often needs to respond very quickly to changes in their visual environment and thus cannot spend too much time collecting light. Further the insect eyes have very small apertures in the ommatidium, resulting in poor contrast vision. To determine the sensitivity and thus the contrast, it is important to know the number of photons available to receptors under various illumination conditions. The number of photons reaching the light sensor in the ommatidium is given by

$$N = 0.621 I D^2 \frac{d^2}{f^2} \tag{2.8}$$

where I is the incident light intensity, D is the ommatidium facet diameter, d is the visual pigment (rhabdom) diameter and f is the focal length of the ommatidium. Usually the time for which the insects eye collects light is in the range of milliseconds to form an image. The absolute threshold for vision is about 1 photon every 40 minutes, which is same for insects and human eyes. The sensitivity of an imaging system is given by the intensity in the image plane of the lens multiplied by the area of the receptor with an absorption coefficient k, as shown in equation (2.9)

$$S = \frac{\pi}{4(F/\#)^2} \frac{\pi}{4} d^2 k = \frac{\pi}{4}^2 \frac{D}{f}^2 (1 - e^{-kl}) \tag{2.9}$$

where k is the natural extinction coefficient of the photo pigment in the receptor of length l, *F/#* is the F-number of an optical system. Equation (2.9) can also be written as equation (2.10), which shows the trade-off between the sensitivity and resolution of the eye

$$S = \frac{\pi}{4}^2 D^2 \Delta\delta^2 (1 - e^{-kl}) \tag{2.10}$$

where $\Delta\delta = d/f$. If enough light is available then $\Delta\delta$ can be small, allowing a resolution close to the diffraction limit. $\Delta\delta$ increases for low light condition and thus degrades the resolution. The compound eyes have a F/# similar to the single aperture eye to have the same sensitivity. In superposition eyes, the F/# is much smaller, making them more sensitive to light compared to apposition eyes.

Human eyes are also best suited for day light and they are not much sensitive in moonlight or starlight. In low light often the human eyes combine visual signals neurologically. In superposition eyes, the sensitivity can be increased by either optically summing the available light from neighboring ommatidia or by summing the electrical signal neurologically. In the former, the light available in different ommatidia are focused onto a single point while in the later the transduced signal is being added. This increases the sensitivity in the dim light. For example using spatiotemporal summation, the locust can extend its vision to light intensities more than 100,000 times dimmer than if it relies on its optics alone. However the visual acuity or spatial resolution is decreased. The summation strategies also reduces the temporal resolution. The spatial resolution is also a function of number of images formed and is shown in figure 2.6.

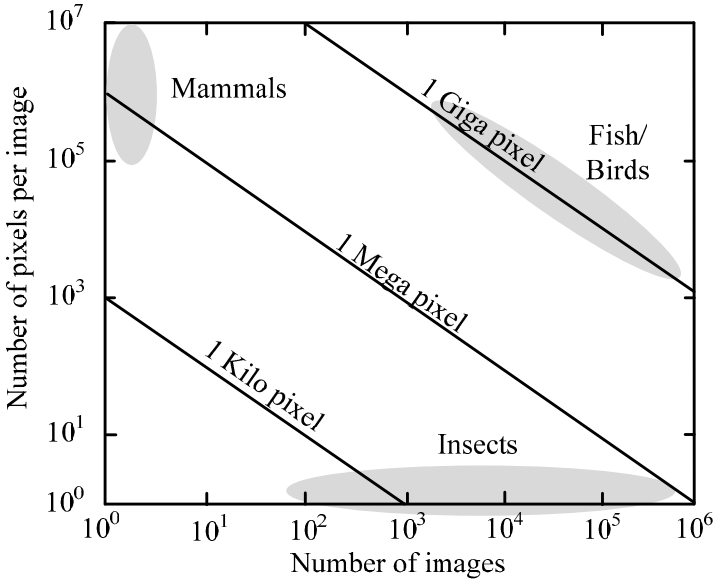


Fig. 2.6 Trade-off curve between spatial resolution and number of images formed

Angular Sensitivity Function

The angular response of an ommatidium to an object placed at an angular distance of ϕ, is given by angular sensitivity function (ASF).

$$ASF(\phi) = e^{\left[-4ln2\left(\frac{\phi}{\Delta\varphi}\right)^2\right]} \tag{2.11}$$

where $\Delta\varphi$ is the full width of the function at 50% sensitivity and is also called as the acceptance angle. The direction of light contains an enormous amount of information about the surrounding world.

Acceptance Angle

The size of the acceptance angle $\Delta\varphi$ defines, the minimum distance between two bright point sources which can still be resolved by the optics. The acceptance angle can be approximated as

$$\Delta\varphi = \left[\left(\frac{d}{f}\right)^2 + \left(\frac{\lambda}{D}\right)^2\right]^{\frac{1}{2}} \tag{2.12}$$

where λ is the wavelength of light. The acceptance angle of an ommatidia has two components, one an optical component, represented by the point-spread function of the lens, λ/D. This is the full-width-half maximum of the Gaussian approximation of the airy function describing the blur caused by diffraction. The point-spread function defines the distribution of light in the image of a point source. The other component is the rhabdom acceptance angle, *d/f*.

The angular resolution of the optical channels is given by $(1/\Delta\varphi)$. The ommatidia provides a large focal depth because of an extremely short focal length. This causes the image plane to be always located in the focal plane in an ommatidium, independent of the object distance. The angular resolution is thus constant over a large range of object distances. However the spatial resolution scales linearly with the object distance.

Modulation Transfer Function (MTF)

Modulation Transfer Function (MTF) is often used to describe the quality of the eye's optics. MTF is the magnitude response of the imaging system to sinusoids of different spatial frequencies. The response is described in terms of the modulation depth, which is a measure of visibility or contrast. Thus MTF can be defined as the ratio of the image and object contrast for grating of all spatial frequencies [5]. The MTF is mathematically related to point spread function (PSF) of the optical system. The MTF is obtained by the Fourier transform of the ASF in equation (2.11), normalized to unity at zero angular frequency v [8]

$$MTF(v) = e^{\left[-\frac{\pi^2}{4ln2}(v\Delta\varphi)^2\right]} \tag{2.13}$$

The finest resolvable sinusoidal pattern is defined by the optical cut-off MTF of the lens of a single ommatidium. The resolution of the compound eyes are very low because of the small diameter of the ommatidium.

Field of View (FOV)

Each optical channel in the natural compound focuses the light coming from the object points lying on the channels optical axis. Due to this bending, each ommatidium points towards a different angular direction. Therefore these compound eyes exhibit a very large field of view *(FOV)* while the single channels are working

on-axis and are not suffering from off-axis aberrations. The field of view of an eye is defined as the angle in object space, over which objects are observable by the eye.

The *FOV* is determined using the equation

$$FOV =\sim 2arctan\left(\frac{D}{2f}\right) \tag{2.14}$$

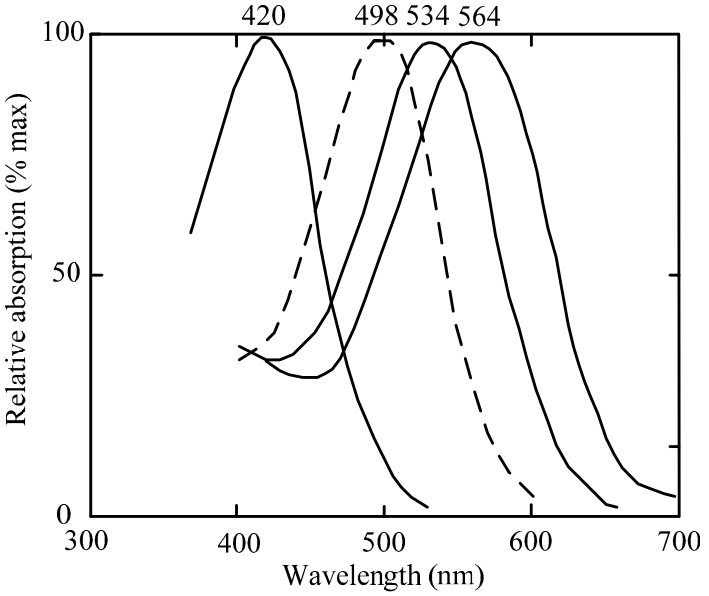
The compound eyes do not suffer from the $\cos^4$-law, which is responsible for the decrease in relative illumination with increase in field of view. The compound eyes are thus optimum arrangement for a high *FOV* and results in high binocular field providing stereoscopic vision.

Colour Vision and Spectral Sensitivity

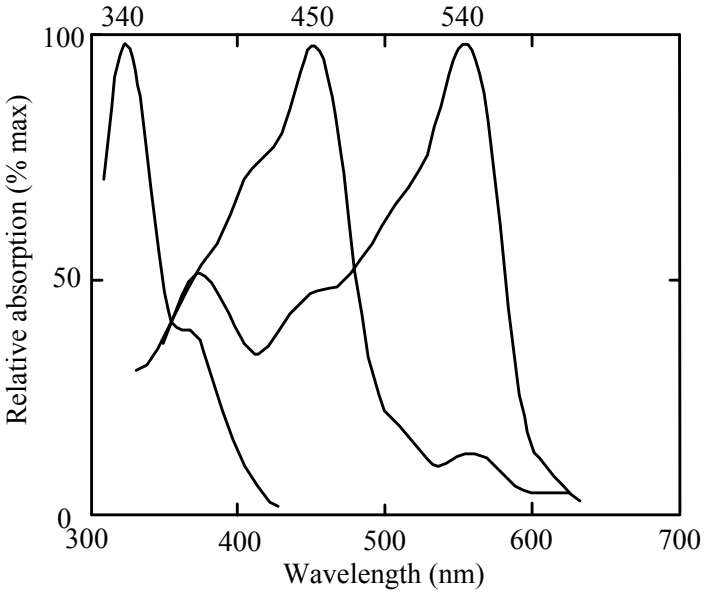
Colour vision is an advanced vision and not many animals can see colour, including cats and dogs, which only see in grey, black and white. The rhodopsins come in several types, with its characteristic behaviour to a certain colour or wavelength of light. The humans usually have three types of colour sensors in their retina that respond to blue light, green light or the red light. Thus the human vision is often called trichromatic (three colour) colour vision. The whole spectrum of hues that human can see is generated by mixing the three primary colours: red, green and blue. A colour sensitivity of eyes can be described by three parameters: hue, saturation and brightness. The hue defines the purity of the colour. Saturation describes the vividness or the intensity of the colour. As saturation increases, the colours appears more pure while they appear washed-out if the saturation decreases. Brightness is the relative intensity of the energy output of the visible light source.

The spectral response of a human eye and a bee's eye is shown in figure 2.7. Most of the insects have trichromatic colour vision, similar to humans, however there visible spectrum is shifted towards shorter wavelengths. The bees and mostly all insects cannot see red light well, however their eyes are sensitive to ultraviolet light. The spectral response of an insects eye depends on whether the rhabdom are open, fused or tiered [9]. Infused rhabdoms, the spectrally different photo-pigments act as lateral filters for one another and thus the spectral response is similar to the visual pigment absorption spectrum. In open rhabdoms, the filtering effects is absent and thus the spectral response is broader than the visual pigment absorption spectrum. In tiered rhabdoms, the distal cells filter the light reaching the proximal photo-receptors and thus the spectral response is narrowed [9]. Some birds and gold fish are tetrachromatic (four primary colours) and thus can see more hues than humans.

In many insects the sensitivity to ultraviolet light is much higher than to blue or green light. For example in drone bees, the sensitivity of a large number of photo receptors increases to the shortest ultraviolet wavelengths measured down to 318nm. The higher sensitivity to ultraviolet light in many insect's eyes help insects to navigate. Sunlight forms a pattern of polarized ultraviolet light in the sky, the insects are able to detect these polarization pattern and able to judge the position of the sun. Thus they are able to navigate using sun as a compass.



(a)



(b)

Fig. 2.7 Spectral response of (a) human eye (b) bee eye

The insects eyes also show adaptation to their visual environment. For example in apposition eyes, in dark conditions the rhabdom molecules move upward allowing light from the neighbouring ommatidia to fall on it. Effectively the optical apposition eyes tends to function as optical superposition eyes. In general the light flux is regulated in apposition eyes by a radial movement of visual pigments in the retinula cells and by changing of the refractive index of waveguide surrounding the rhabdom. Human also show dark adaption. When the room is made suddenly dark, for the first few second nothing is visible but after some time, things will be dimly visible.

The restricted integration time over which the photons are integrated by the visual molecules, introduces motion blur when the eye and the surroundings move relative to each other. The image starts to lose contrast at high spatial frequencies when the relative motion exceeds one acceptance angle per receptor integration time [5]. The conventional television screen used in homes refreshes every 25 or 30 times a seconds. Many electrical lights also flicker at a rate of 100 or 120 times a second. However these flickering rate is so high that its not visible to the human eye. Thus for any visual stimulus whose flicker frequency is higher than the visual system, the stimulus appears continuous. The flicker frequency of human eye is only 15-20 times a second while the flicker frequency of insects is very high, for example it is 300 for honeybee. Thus the temporal resolution of the insects eye is much higher than the human eye. In apposition eyes, wider facets (D) and wider rhabdoms (d) increases the visual sensitivity by 1-2 log units.

A comparison of different eyes and related parameters are shown in table 2.1.

Table 2.1 Comparison of single aperture, apposition and superposition eyes [8]

Parameter	Human	Bee	Moth
Eye type	Single aperture eyes	Apposition	Superposition
Light habitat	Diurnal	Diurnal	Nocturnal
Lens diameter (mm)	7	0.025	0.4
Focal length (mm)	23	0.06	0.17
Receptor diameter (μ m)	2	1.5	8
F-number (F/#)	3.3	2.4	0.4
Sensitivity (μm^2=10^{-12}W/(W/m^2))	0.23	0.24	218
Acceptance angle (°)	0.007	1.9	$<$13
Interreceptor angle d/f (°)	0.005	-	3
Interommatidial angle (°)	-	0.95	-
Resolution (LP/°)	100	0.52	$>$ 0.08

2.3 Polarization Vision in Compound Eyes

The most important use of the sophisticated eyes of many higher vertebrates and many invertebrates are in image formation and motion detection. They are however

also able to detect the other properties of light such as the wavelength (colour) or the vector of polarization. The sensitivity to polarization has been shown to be useful in a number of visual functions including increase in contrast, orientation, navigation, prey detection, predator avoidance, object recognition, camouflage breaking and intra-species signal detection and discrimination.

The compound eyes of insects are formed by hundreds or thousands of simple eyes called ommatidia. Each of these sub-units has its own lens, crystalline cone, and several long visual cells arranged in a star pattern as described in section 2.1. The light sensitive parts of the visual cells are the microvilli which are an array of tube-like membranes where the pigment rhodopsin is located. In all animals, invertebrates and vertebrates, the visual pigment rhodopsin is present in the photoreceptor membrane of the visual cells. These photoreceptors are preferentially sensitive to a specific *e*-vector direction due to their strong dipole moment. They are maximally excited when the electrical vector (*e*-vector) axis is parallel to the dipole axis. In vertebrates, these photoreceptors are oriented perpendicular to the paths of the incoming light rays, presenting a random array and thus are typically insensitive to polarized light. In invertebrates, the photoreceptors are aligned parallel to the microvillar axis. Thus if all microvilli of a single photoreceptor cell are parallel, the cell will respond most strongly to incoming polarized light with its *e*-vector aligned parallel to the microvillus, thus exhibiting polarization sensitive behavior [10].

The detection of polarization in the eyes of insects is mediated through the ommatidia present in the dorsal rim area *(DRA)* of the compound eye. In each ommatidium, the photoreceptors come in two sets that have their microvilli oriented at about 90° to each other. The receptors are thus tuned to mutually orthogonal *e*-vectors promoting polarization antagonism [11]. The high polarization sensitivity of the *DRA* is attributed to the microvilli being aligned along the rhabdomere. The polarization sensitivity of the ommatidia outside the *DRA* of the compound eye is weak, because of the misalignment of the microvilli [11]. The variations in the ommatidia in the *DRA* and outside of the compound eye of the field cricket are shown in figure 2.8.

Polarization of light has been used as visual signals by many insects. The insects have build receptor systems capable of analyzing the polarization information of incoming light. The polarized light is abundant in nature, commonly being produced by scattering or reflection. The mechanisms by which an unpolarized light becomes polarized is discussed in chapter 3. Polarization based signaling are quite different from conventional colour based or other forms of visual signals. The polarized based signalling do not vary with habitat depth unlike signaling based on spectral signatures.

Some animals like Octopus are able to discriminate polarization variation in a single object and thus be able to segregate the objects polarization features [12]. This is similar to colour vision wherein similar bright light reflected varies in its spectral content. The sea animals use polarization of light more often as the spectral irradiance underwater varies strongly with depth, but the polarization is very stable and thus the polarization signal are consistent. The stomatopod species uses

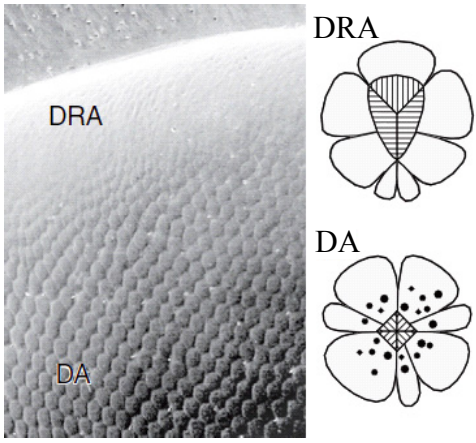


Fig. 2.8 Cross section of ommatidia in the DRA and outside DRA (DA) regions of the compound eye of the field cricket in the dorsal most part of the compound eye [11]

reflection of strongly polarized light for interspecific communication. Polarization based signaling have special properties that make them particularly suitable for use in environments where the natural polarization field is weak, stable, or highly predictable, as they will stand out against natural backgrounds [13].

Polarization of light is also used by insects for contrast enhancement. The polarization sensitivity can be used to increase the contrast by [14]

- detecting the changes in the degree of polarization of the background light transmitted through birefringent tissues
- detecting the change of angle of the polarization of light after being transmitted through otherwise transparent tissue
- subtracting the polarized path radiance and background radiance and
- subtracting the polarized path and background radiance and compensating for the attenuation of the directly transmitted light.

2.4 Artificial Compound Eye

Early research in machine vision was aimed at developing systems for general purpose vision. The algorithms implemented were aimed at extracting as much information as possible from a single scene, or from a small number of scenes captured at fixed locations. Over the last few years the research has focused more towards active vision systems where there is an interaction between the viewing device and the external world. Traditionally, problems in machine vision have been tackled through reasoning that is based primarily on geometry, mathematics and logic, with little consideration of how these problems might be solved by natural visual systems. This approach produces solutions that are often computationally complex, sensitive to noise, and sometimes too general purpose to be really effective.

There is a growing trend towards exploring strategies for artificial vision that take advantage of the biological systems wherein there is more interaction between the viewing device and the external world.

The artificial compound eye has a number of advantages over ordinary single chamber optical system. Since it is a single sensor with multiple lenses each eyelet lens contributes its own sub-image from different directions to a large field-of-view, which facilitates sub-imaging and parallel image processing. Research on artificial compound eyes falls into two categories: (a) multiple camera systems wherein the scene is electronically distributed across multiple lenses using multiple detector arrays or (b) single camera system wherein the scene is optically distributed across multiple lenses using a single detector array. Most of the compound eye literature falls into the first category, although the second category has been catching up in the last few years [15].

Leonardo da Vinci [16] stated that "every body in the light and shade fills the surrounding air with infinite images of itself; and these, by infinite pyramids defused in the air, represent this body throughout space and on every side" as shown in figure 2.9. A plenoptic function is used for the visual representation of the scene observed. This function changes differentially over time. Each of the "pyramids" in figure 2.9 denotes a subset of the plenoptic function at a given time which can be captured by an imaging device. The plenoptic function is described in more detail is chapter 7.

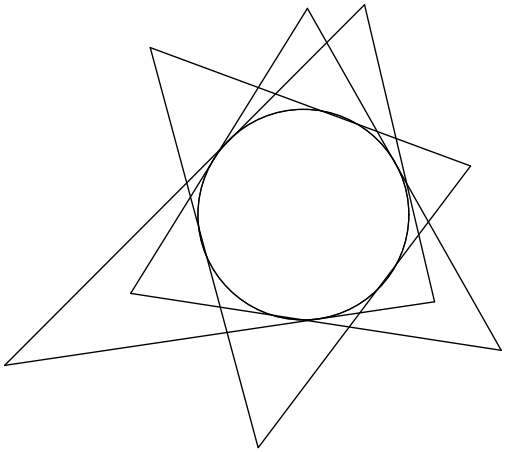


Fig. 2.9 Diagram of optical "pyramids" from Leonardo's notebook [17]

In the conventional imaging approach using a single pinhole camera, only one of the pyramids in its field of view is visible and thus forms a single image. A single-axis lens focuses an image on the focal plane of the image sensor, the resolution of which is determined by the size of the optical window. Through this arrangement,

cross talk, i.e. the coupling of one signal into the adjacent channel is negligible even for significant aberrations or misalignment errors and a high resolution is obtained over a continuous area. However conventional imaging systems are not able to convey the shape and size of the object of interest.

In the multichannel aperture imaging system, the plenoptic function from many different viewing points are captured. Multichannel electro-optical systems are superior in detecting motion and determining the shape of the object in motion. To determine the structure of an object from motion two conditions need to be satisfied: high field of view and linearity of estimation [18]. In single pinhole cameras, one ray from each point in space is captured and from this ray at different times the structure of motion must be estimated. This makes the estimation of the viewing geometry a non-linear problem. The additional information in the multichannel imaging system (multiple rays from the same scene point) makes estimation of the viewing geometry linear.

The model of an artificial compound eye vision system based on multiple camera system is shown in figure 2.10.

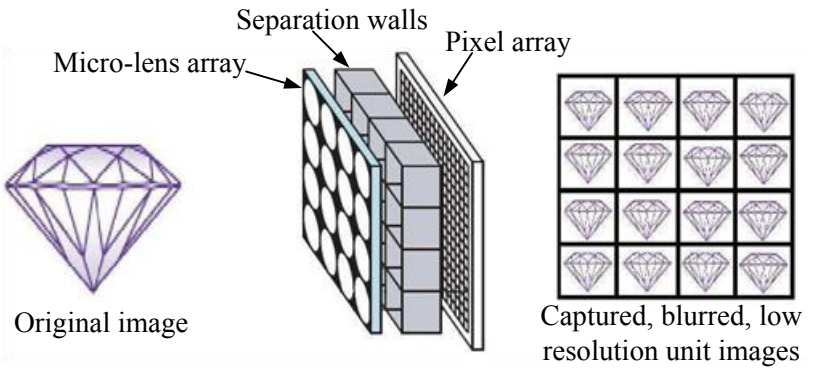


Fig. 2.10 The TOMBO (Thin Observation Module by Bound Optics) architecture [19]

Tanida et al. in 2000 developed an imaging system based on compound eye, consisting of a micro-lens array with planar structure, an optical separate layer and a photodetector array called TOMBO (an acronym for Thin Observation Module by Bound Optics). A micro-lens and its array of photosensitive cells together form an imaging unit. Each micro-lens of the array imaged the object independently. The optical separate layers were used to prevent optical interference between the optical channels. The TOMBO sensor captures multiple low resolution images at the same time and the output is a mosaic of low resolution images from individual optical channel. The final image is constructed from the multiple images acquired though each micro-lens using digital processing. The process of obtaining high resolution images from low resolution images involved complex image processing.

A theoretical model of a multichannel or multiple apertures imaging system similar to the compound eye of insects is shown in figure 2.11. In the multiple lens system each lens, helps to form a sub image of the object seen with a different angle in the field of view.

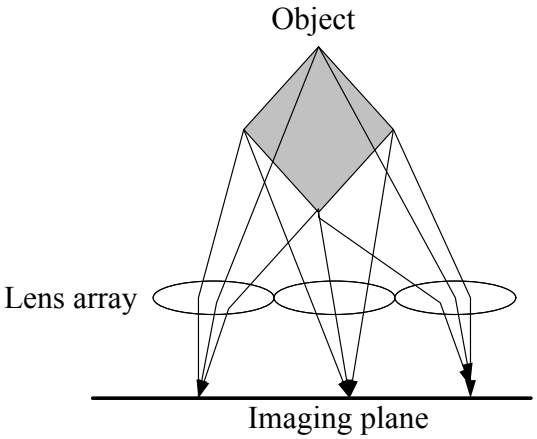


Fig. 2.11 The multiple lenses capture each pyramid of the same object

If each lens in the lens array has n sub-pixels along one dimension then the spatial resolution of the system is reduced by a factor of n, producing low resolution sub-images at each lens [17]. This photographic technique has already been described as "Integral Photography" (because the resulting integral image is the sum of individual image) by G. Lipmann and H. E. Ives. Many possible implementation of such a multiple lens imaging system can be found in literature today [17], [20], and [21]. Due to the reduced imaging field for each micro-lens in multichannel imaging, the requirements on the micro-lenses are relatively moderate, although high resolution is still required. Implementation of the multichannel imaging system can be significantly cheaper than the conventional approach of using multiple single-axis cameras. The problems of multichannel imaging are related to crosstalk and a limited interconnection distance.

Duparre et al. introduced another artificial compound eye structure using planar micro-lens array called APCO (Artificial Apposition Compound Eye). Different from the Tanida's model, each micro-lens in the system imaged only a small portion of the object in their field of view. The final image is formed by splicing the partial images directly. The model of an artificial visual system based on the single camera system with multiple lenses/aperture to transfer the image through separated optical channels [22] is shown in figure 2.12. The artificial apposition compound eye consists of a micro-lens array positioned on a substrate, preferably with optical isolation of the channels, and an optoelectronic detector array of different pitch in the micro-lenses' focal plane. The pitch difference enables the different viewing

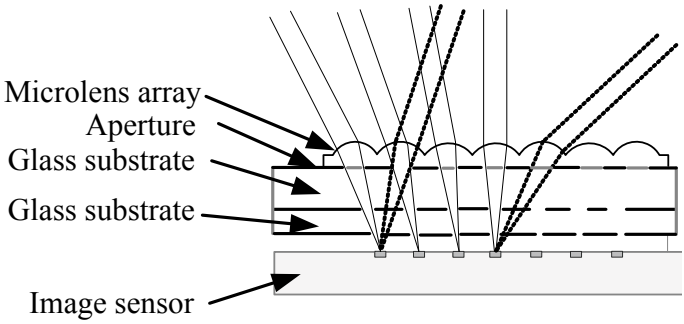


Fig. 2.12 Structure of an artificial compound eye

directions of each optical channel. Each channel's optical axis points in a different direction with the optical axes of the channels directed outward if the pitch of the receptor array is smaller than that of the micro-lens array.

The micro-lens array was fabricated on a planar glass substrate by photo-resist reflow method. A titanium pinhole array was fabricated on the back of the substrate by photo-lithography and wet etching. The micro-lens array and the pinhole array were aligned, such that each pair of the micro-lens and the pinhole constitutes an optical channel. The pitch of the pinhole array is different from the pitch of the micro-lens array, for each channel to have a different view angle. If the pitch of the micro-lens array is larger than the receptor array, an upright image is obtained and an inverted image is obtained if the pitch of the micro-lens array is smaller than the receptor array. The optical axis of each channel points into a different direction of object space. The focal depth of the APCO system is large, due to the extremely short focal length of the micro-lenses.

The group of Duparre also developed a compound eye imaging system with a chirped micro-lens array. In the chirped micro-lens array, the optical axis direction, aperture size and shape of the micro-lenses are determined according to the view angle of the optical channel. The chirped micro-lens reduces the aberrations significantly and improves the image quality. Compound eye based imaging system using spherical structure has also been reported [23]. Here the micro-lens and the pinhole array are fabricated on two concentric spherical bulk lenses respectively. This design provides for a higher field of view compared to the artificial compound eyes with planar structure.

2.4.1 Design Consideration of an Artificial Compound Eye

A major difference between the insect facet eyes and the artificial eyes is the shape of the detector which is flat in the artificial case and curved for the insect eyes. The arrangement of the micro-lens layer and the detector layer in a natural compound eye is shown in figure 2.5. The angular distance between the two micro-optical channels, $\Delta\phi$, determines the visual acuity or the resolution and is called the interommatidial

angle. The acceptance angle, $\Delta\varphi$, is given by the ratio of *d/f* and the diffraction effects of the micro-lens aperture by λ/D. Each optical channel focuses the light coming from the object points lying on the channels' optical axis. Due to this bending, each ommatidium points towards a different angular direction. Therefore these compound eyes exhibit a very large field of view *(FOV)* while the single channels are working on-axis and are not suffering from off-axis aberrations [24]. The field of view of an eye is defined as the angle in object space, over which objects are observable by the eye. From a technological point of view it is more difficult to work with a curved CMOS detector than with a flat one, although progress has been made in this field [25].

The major design parameters for artificial compound eye are the angular resolution, sensitivity and field of view. If the detector resolution is ignored and the effects of aberrations can be minimized, then the resolution is limited by the diffraction limitation of the lens. For an imaging system which is diffraction limited, the minimum spot size possible is 0.61λ (λ is the wavelength of the incident light), assuming an numerical aperture of 1. The numerical aperture is the range of angles over which the optical system can accept light. The number of resolution points across an image can be determined from the optical and geometrical parameters using an ideal thin lens as shown in figure 2.13.

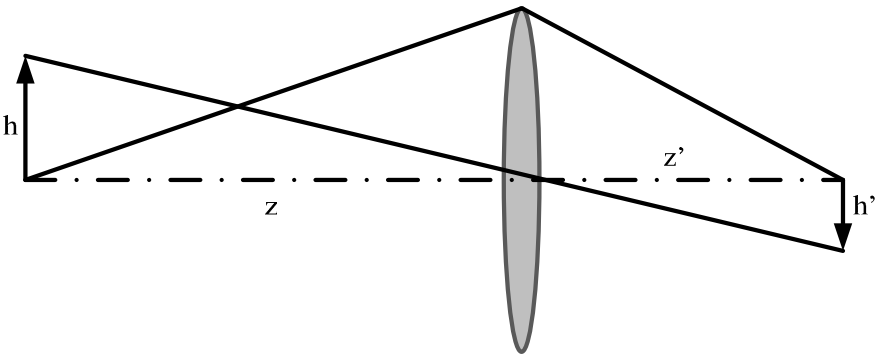


Fig. 2.13 Thin lens (first order) description of imaging system

The number of resolution elements, *N*, can be determined using the equation

$$N = \frac{h'}{w} = \frac{z'h}{0.61\lambda z} \tag{2.15}$$

where w is the diffraction limited spot radius, λ is the wavelength of the incident light, h is the height of the object, h' is the height of the image, z is the distance between the lens and the object and z' is the the distance between the lens and the image.

The angular sensitivity function is used to characterize the performance of a single artificial micro-optical channel. It describes the amount of light flux, radiated from an object is received by the photoreceptor for varying angular distance of the object from the optical axis of the micro-optical channel. The angular sensitivity function is obtained by convoluting the Point Spread Function *(PSF)* of the micro-lens with the pinhole. The micro-lens *PSF* at the micro-lens aperture for a certain field angle is multiplied by the transmission function of the pinhole.

For the APCO system, if the pinhole diameter is large, the sensitivity increases but the resolution decreases due to the overlap of different micro-optical channels field of view. The resolution of a planar artificial compound eye depends on the pitch difference between the micro-lens array and the pinhole array [8].

$$\Delta\phi = arctan\frac{\Delta p}{f} \tag{2.16}$$

where Δp is the pitch difference. The acceptance angle, which determines the smallest resolvable object period, is given by the *PSF* of the micro-lens and the transmission profile of the pinhole array.

$$\Delta\varphi \propto \frac{t_{pinehole} x PSF}{f} \tag{2.17}$$

The maximum resolvable line pairs over the *FOV* is determined by the acceptance angle and is given by

$$\chi_{LP} = \frac{FOV}{\Delta\varphi} \tag{2.18}$$

The *FOV* for a given acceptance angle is desired to be large to be able to obtain large number of image pixels. The *FOV* of an imaging system and the corresponding image size are related as shown in figure 2.14, and must be matched to the digital sensor array size. The required magnification is defined by the focal length which is given by

$$f = \frac{a_D}{2tan\alpha} \tag{2.19}$$

where a_D is the detector diagonal size and α the size of half the diagonal *FOV*. The pixel sizes are determined by the sensitivity and diffraction effects desired.

An important parameter to evaluate the optical quality of a micro-lens is the F-number $(F/\#)$. The F-number of a micro-lens is defined as

$$F - number(F/\#) = \frac{f}{D} \tag{2.20}$$

where f and D are the focal length and diameter of the micro-lens or aperture size, respectively. The F-number expresses the angular aperture of a lens. The numerical

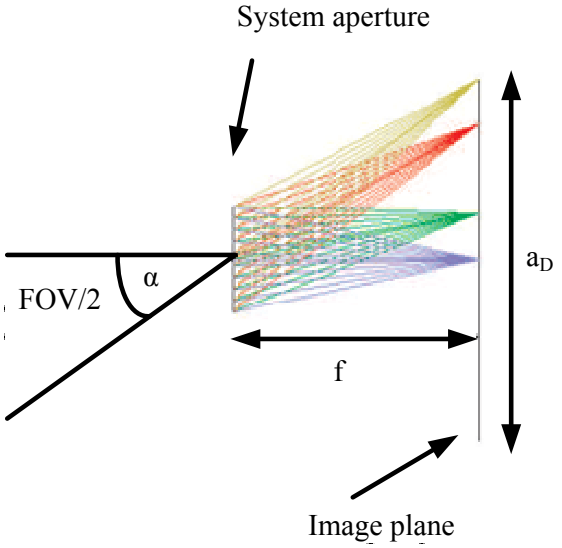


Fig. 2.14 The FOV of the imaging system and the image size [8]

aperture *(NA)* of an optical system characterizes the range over which light can enter into an imaging system. The numerical aperture determines the finest details that can be resolved by the optical system. A lens with larger numerical aperture will be able to visualize fine details than a lens with smaller numerical aperture. The numerical aperture in terms of F-number is given by

$$NA = sin\left(arctan\frac{1}{2F/\#}\right) = sin\left(arctan\frac{D}{2f}\right) \tag{2.21}$$

The size of the lens affects the image formation and thus is very important. Small lenses transfer fine details of the object just like large lenses do, but are limited by the number of details that they can send as shown in figure 2.15. Each channel transfer the information as per its field of view, and thus it makes sense to have more number of channels in parallel to capture larger extent of the visual environment.

The number of information in the image plane of a large lens is very large because of the small airy disk due to lower diffraction and the large focal length. As the size of the lens decreases while the size of the image point remains the same, the spatial information collected by the small lens is low because of the small image field. By applying an array of micro-lens as shown in figure 2.15, the spatial information capacity can be increased as the image filed in transferred in parallel and thus has the same capacity as the large lens. In micro-optics, lens abberation is not critical, the diffraction limit is the most important parameter which defines the image resolution. The diffraction limit can be overcomed by using multiple lens array to collect the information from the visual field.

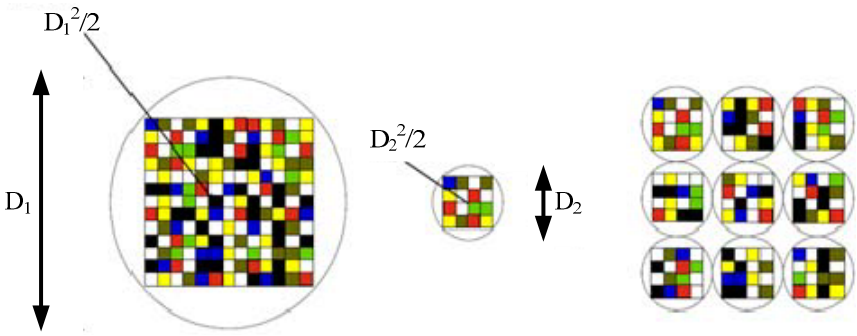


Fig. 2.15 The relationship between the lens size and the image size [26]

The sensitivity or light collection ability of the optical system is also an important property. The sensitivity of a single optical channel can be written as

$$\frac{P_I}{I_O} = \tau N A^2 \frac{\pi d^2}{4} \tag{2.22}$$

where P_I is the incident optical power through the pinhole, I_O is the irradiance in the object plane, τ is the transmission of the optical system, d is the pinhole diameter and *NA* is the numerical aperture of the micro-lens.

$$NA \approx \frac{D}{2f} \tag{2.23}$$

The size of the acceptance angle determines the trade-off between sensitivity and resolution in an artificial compound eye. A major problem with artificial compound eyes is the inverse relationship between resolution and sensitivity. The optical power increases with the diameter of the pinhole array however the resolution decreases. The superposition eyes solves this problem by optically summing the light from various optical channels. The micro-lens diameter D influences the sensitivity through the micro-lens numerical aperture. The numerical aperture expresses the ability of a lens to resolve fine details in the object being imaged. The pinhole diameter d needs to be very small to achieve high resolution, this however leads to a lower sensitivity. In neural superposition eye, each object point is imaged by multiple pixels independently. The information from each of these pixels is then superimposed to form the final image, thus increasing the sensitivity of the eye.

2.4.2 *Design of Artificial Compound Eye*

The two major design elements of an artificial optical visual system as shown in figure 2.11 are the micro-lens array and photoreceptors or the image sensor. As stated in chapter 1, the performance of the eyes depend on the angular spacing of the receptors, which determines how finely an image can be resolved. The interommatidial angle in insects is typically 1° (0.0175 rad) and thus an angular

resolution of near 1° is desired for the micro-lens array. The angular resolution is defined as the smallest angular difference between two incident beams which results in non-overlapping point spread functions *(PSFs)*. *PSF* is defined as the distribution of the relative intensity over relative distance. Usually for every *FOV*, diffraction is more important than geometrical abberations, however optical simulations showed that for a *FOV* of near 25°, geometrical abberation becomes more important than the diffraction. Thus for *FOV* greater then 25°, the angular resolution worsens. To have the desired angular resolution of 1°, individual micro-lens with a *FOV* of 25° was designed. An array of such individual micro-lenses was used to enhance the total field of view of the micro-optical system. The design of the micro-lens array is discussed in section 2.5.

Due to robustness and cost effectiveness a flat CMOS detector was selected. The optical simulations showed that in flat facets (detectors), due to the diffraction issues reducing the facet dimensions below 25μm doesn't help to obtain a better focusing. Furthermore the dimensions of the photodetector should not be smaller than 10μm to receive all the available photon flux. In curved facets however a better focussing is obtained for facets larger than 25μm. To utilize the entire area of 5mm x 5mm available in a multiple wafer run *(MPW)* from Europractice, an imaging array of 128 x 128 was selected with a pixel pitch of 25μm and a photodiode area of 10μm x 10μm. An image sensor with the these specifications was designed and is described in detail in chapter 3.

2.5 Micro-optics Design

In this section, the micro-optical lens designed by the TONA group of Vrije Universiteit Brussels *(VUB)* as part of a collaborative project is described briefly. The system uses a micro-lens array to form multiple images, which are captured on a photo-detector array. This emulates the compound eyes of insects. The goal is to design a multichannel imaging system with a wide field of view, in this case near 124° with an angular resolution of less than 1° and also to find a good trade-off between the image resolution and sensitivity. A brief overview of the micro-lens design is presented in this section.

The micro-lens array can be fabricated using techniques like, ion exchange method, photosensitive method, thermal reflow method, chemical vapor deposition, hydrophobic method, focused ion beam method, hot embossing, micro-machining etc. The thermal reflow method, because of its low cost and simple fabrication method has become the most popular micro-lens array fabrication technique. The thermal reflow method is shown in figure 2.16. In the first step of the thermal reflow method, a photo-layer is spin-coated onto the substrate and and it is then exposed by ultraviolet light. The photo-resist layer is then developed and cylindrical isolated array pattern is formed. These cylindrical structures are then heated till they melt and turn into spherical structures.

The contact angle, α, shown in figure 2.16, is a key parameter which determines the quality of the spherical micro-lens produced using thermal reflow method. In figure 2.16, the contact angle is given by

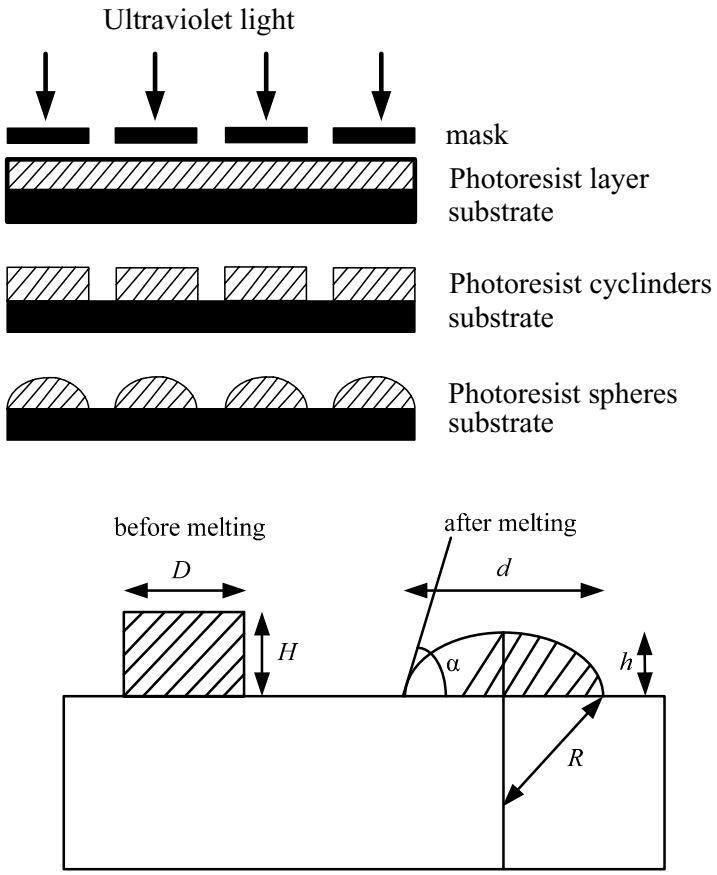


Fig. 2.16 Thermal reflow process for fabrication of lens

$$cos\alpha = \frac{R-h}{R} \tag{2.24}$$

where R is the radius of curvature and h is the height of the micro-lens. The radius of curvature of the compound eye is related to the focal length as

$$R = f(n-1) \tag{2.25}$$

where f is the focal length of the micro-lens and n is the refractive index. Further from geometrical considerations in figure 2.16,

$$sin\alpha = \frac{d}{2R} \tag{2.26}$$

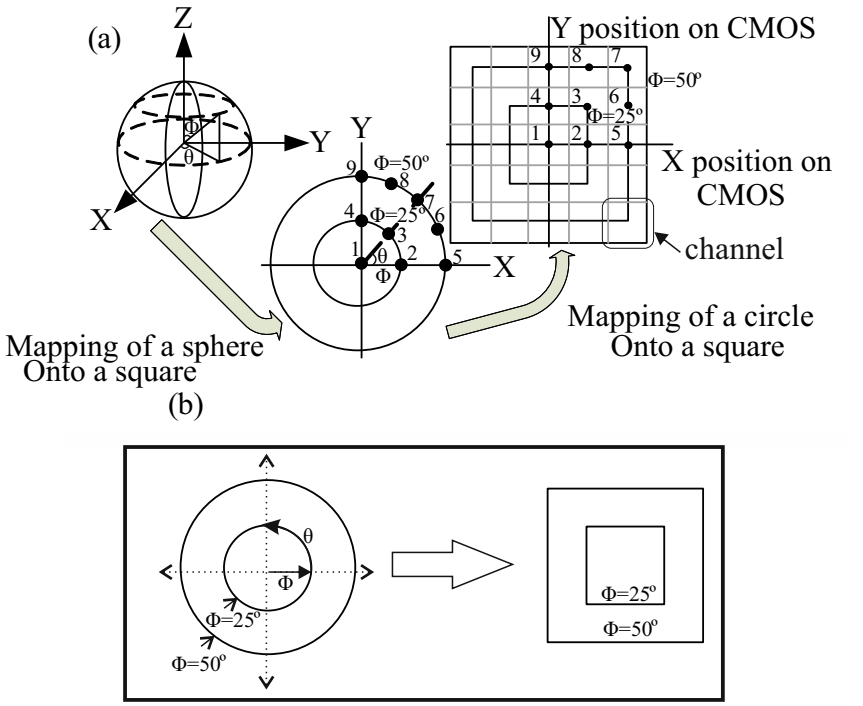


Fig. 2.17 Angle to position transformation: (a) Mapping of a sphere onto a square. (b) Angles phi (φ) and theta (θ) for the center directions of angles of incidence for the first nine channels.

Using the above equations, the relationship between F-number and the contact angle can be determined as

$$F - number = \frac{1}{2(n-1)sin\alpha} \tag{2.27}$$

The radius of curvature *(R)* of the eye determines the maximum achievable visual acuity as discussed in section 1.2. Thus the contact angle will define the F-number of the micro-lens fabricated.

The technological problem of mapping the spherical world onto the planar photodetector surface is solved by dividing the detector plane into 25 zones and an array of 5 x 5 micro-lenses covering all the zones. Each zone has a total *FOV* of near 25°, and capture the light from certain angles with a resolution of 1°, while there are 25 x 25 pixels under each micro-lens.

Figure 2.17 shows the 25 separate channels designed to map the spherical word onto the planar surface. Each point in the far field of the image sensor can be described by two spherical coordinates: the angles phi (φ) or azimuthal angle and theta (θ) or the elevation angle. The CMOS detector is divided into 25 equal zones

and each zone maps a circle with radius equivalent to the elevation angle φ. A quarter of the total *FOV*, is separated into 9 channels by varying θ in three steps for φ=25°and in five steps for φ=50°as shown in figure 2.17.

To sample a wide field of view *(FOV)* of 124°, an overlap of the different angular regions mapped by the separate optical channels is allowed. Figure 2.18 shows a quarter of this total *FOV*. It is very difficult to avoid overlap when we want to cover the complete *FOV* due to the problem of mapping a spherical angular distribution to a square detector.

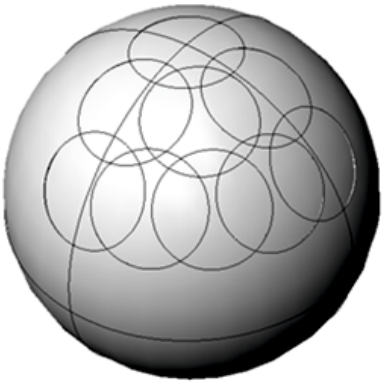


Fig. 2.18 Overlap of the different sub-FOVs

The micro-lens array can either have one lens for one angle, where each lens samples one direction in space or one lens for more angle, where each lens samples more directions in space. Real insect eyes use one or more lenses to sample the light from one direction in space, the focus is still good as the curved facet allows the light to be focused on the center of the detector. For flat facet, this can be achieved by tilting the zones to improve focussing, but fabrication is too difficult and the detection plane should be tilted too. Thus in order to achieve the desired 1° angular resolution, a two lens structure is used as shown in figure 2.19(a). The function of the first lens is to capture the correct part of light from the total *FOV* and to have some focusing towards the second lens. The second lens focuses the light towards a spot on the detector. The scanning electron microscopy *(SEM)* photographs of the top and bottom lenses of the optical channels are shown in the figure 2.19(b), (c). The lenses, substrates and pedestals are fabricated in poly-methyl-methacrylate *(pmma)* material.

Figure 2.20 shows the simulated light spots on the detector. The central spot and the outer spots originating from the incident rays for which the channel was optimized, are shown. The other spots come from incident rays with an angular difference of 1°. It is clear that all the spots are distinguishable on the detector.

A quantitative merit function for the image quality is the contrast of the image, represented by the modulation transfer function *(MTF)* as discussed in

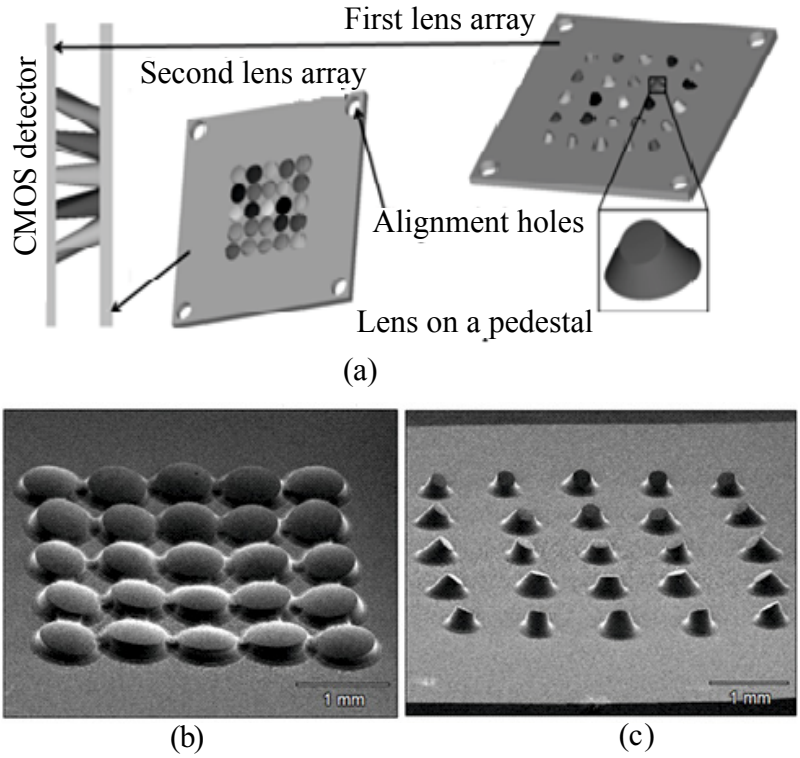


Fig. 2.19 (a) Cross-section of the two lens structure, first lens array, capturing the total FOV, second lens array focusing the light on to the detector with 1°angular resolution (b) SEM photograph of the first lens array (c) SEM photograph of the second lens array [27]

section 1.2. The simulated *MTF* for the described lens array (channels 1-9, as shown in figure 2.17) is shown in figure 2.21.

A maximum contrast of 0.3 line pairs per degree (LP/ °) and an angular resolution of 1° (down to 0.5° for some channels) with a total *FOV* of 124° have been presented [27]. This system is comparable in resolution to others found in literature, however the *FOV* is much better [15], [22].

As discussed, to sample the wide field of view, an overlap of the different angular regions mapped by the separate optical channels is allowed. This means that certain directions will be sampled by more than one pixel of the CMOS detector, introducing distortions in the captured image. However the distortions in the images can be corrected because there is a direct relationship between the angle of incidence and position on the detector. It was not possible to test the micro-lens array with the photodetector array, due to the delay in fabrication of micro-lens array, thus the distortions could not be quantified. The image distortion correction algorithms that can be used to correct the distorted image are described in detail in the following sub-section.

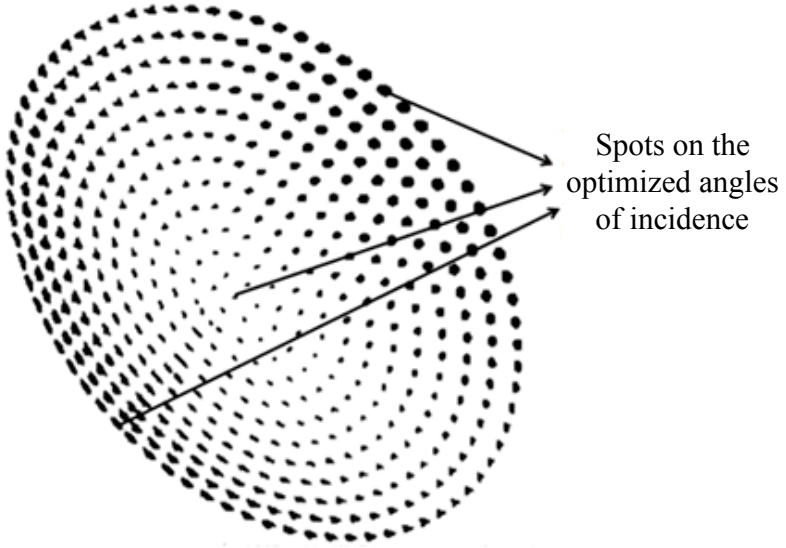


Fig. 2.20 Channel seven: Light spots on the detector. There is an angular resolution of 1°.

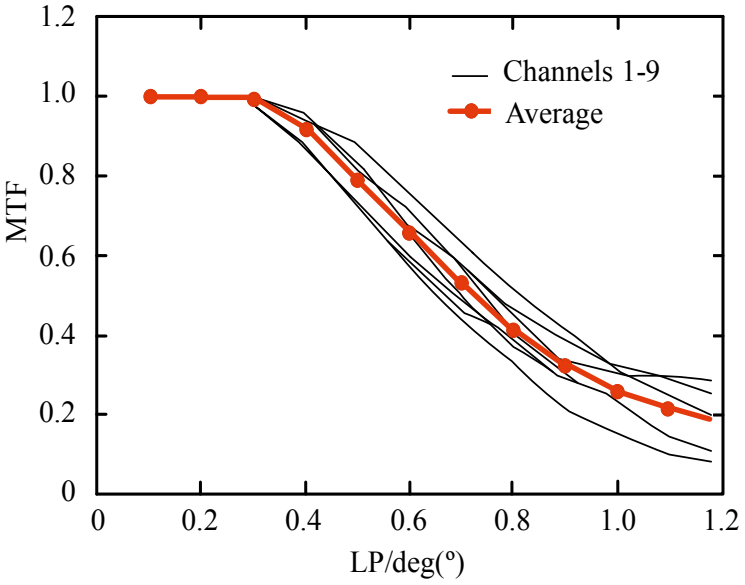


Fig. 2.21 MTF for the different channels of the double tilted lens design

2.5.1 Image Mapping and Distortion Correction

The captured images from the multiple lens optical system are processed to reconstruct the proper image. Each optical channel captures light from a certain *FOV*, but due to the mapping of spherical angular distribution onto a square detector, overlapping of the *FOVs* is not completely avoidable. Thus certain directions will be sampled by more than one pixel of the CMOS detector. This introduces non-linearities or distortion in every zone by the optical system, as shown in figure 2.22. The figure shows the simulated spot resolution on the detector for channels one to nine from the designed 25 channels as described in section 2.3. The spot distribution obtained by the different optical channels are well separated which shows that there are no stray-light. Furthermore the energy decreases towards the outer region of the sub-*FOVs* due to vignetting. In ideal case the spots in each optical channel should be easily distinguishable, however due to the distortion of the optical channel nonlinearities are introduced deforming the shape of the optical channels.

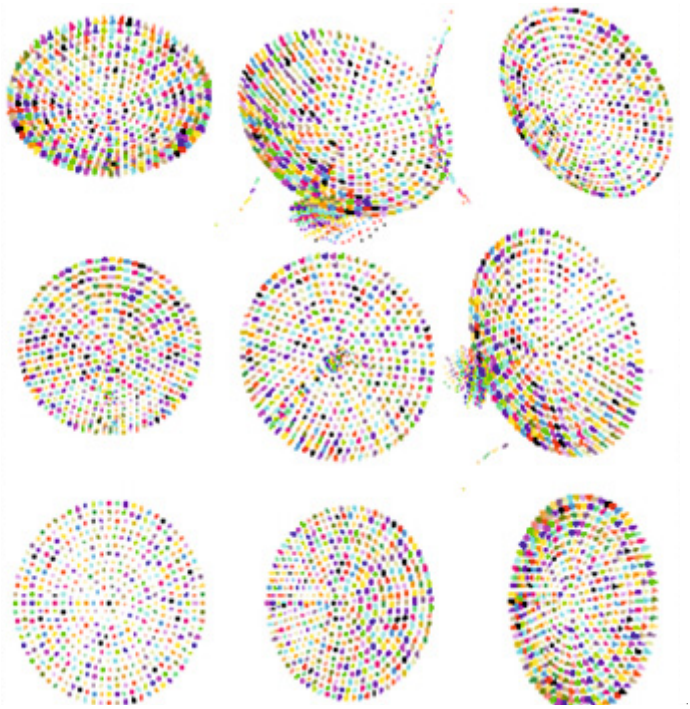


Fig. 2.22 Obtained distorted sub-image from the nine optical channels for the designed lens array

Nonlinearities of an optical system can be largely described by two types of distortion: lens geometric distortion and photometric distortion. The photometric distortion refers to a position-dependent loss of light in the output of the optical

system. It is caused by the varying *PSF* at each pixel due to blurring, scattering of light rays or overlapping of the light rays. This degrades the spatial resolution of the system's final image and it causes gradual fading of the image points near the periphery. The geometric distortion is caused by the variation in the focusing distance. The projection function of a multi lens camera can be observed as that of a projection from a perspective (central) camera, followed by a non parametric displacement of the imaged point in the direction of the distortion center [28]. The displacement of the imaged points induces a non-linear distortion, i.e. straight lines in space are not imaged as straight line unless passing through the center.

Image mapping is limited by the nonlinearities of the optical system. Image mapping is defined as the direct correspondence of the image at the pixel and the viewing direction. A high resolution image can be obtained from the set of low resolution images by pixel rearranging. However it is important to remove the photometric and geometric nonlinearities before trying to rearrange the sub-images pixels on the virtual image.

The photometric distortion can be corrected using the model proposed by [29]. The output image *g(x,y)* can be expressed as

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \tag{2.28}$$

where *f(x,y)* is the original object on the image plane, *n(x,y)* is noise, *h(x,y)* is the system's point spread function *(PSF)* containing the information of the degradation introduced by the system. The noise and the *PSF* for each lens channel is known, thus the original image *f(x,y)* can be easily obtained.

Geometric distortion can be further sub-classified into radial distortion, decentering distortion and thin prism distortion. Radial distortion is caused by imperfect lens shape and manifests itself as radial positional error. Decentering distortion and thin prism distortion are caused by the improper lens and camera assembly, generating both radial and tangential errors in point position. Among the three forms of geometric distortion, radial distortion is of prime concern. In figure 2.22, two types of radial distortion are observed: pincushion and barrel. Pincushion distortion distorts a square object into a pincushion shape while barrel distorts a square object into more of a barrel shaped object. The barrel distortion is more dominant in a lens system, mainly in lenses with short focal length.

Radial distortion is strictly symmetrical about the optical axis of the lens and is governed by the equation [30]

$$r' = r + k_1 r^3 + k_2 r^5 + k_3 r^7 ... \tag{2.29}$$

Equation (2.29) can be safely approximated by using only the first term of the infinite series as

$$r' = r + k_1 r^3 \tag{2.30}$$

where r' is the distorted radius, k is the coefficient of radial distortion and r is the original radius. The sign of k in equation (2.30) affects the type of radial distortion. If k is negative then it is a barrel radial distortion and if positive then it is a pincushion radial distortion.

In order to correct for the radial distortion, the center of the radial distortion and the coefficient of radial distortion are to be computed. The center of the distortion is not the center of the image. However, as the true position is not well constrained and it is necessary to fix it somewhere, the center of the image turns out to be a good approximation [31]. The center of the sub-image needs to be calibrated with high accuracy for two reasons:

a) Each micro-lens has a relatively small field of view with a dimension of around 25x25 pixels in the designed electro-optical system. Thus if the image center calculation varies by approximately five pixels, an error of 20% is introduced in the image reconstruction, which is quite high. A one pixel variation of the center point would reduce the error to 4%.
b) Accurate image center position leads to the prediction of accurate position of the optical axis for each eyelet. The radial distortion is symmetrical around the optical axis, thus the accurate determination of the optical axis would help to eliminate or reduce the radial distortion.

The sub-image center point can be determined using various methods. Most algorithms in literature use the center of the image buffer frame as the center of the sub-image. This method introduces considerable error compared to the field of view. The other common method used in the literature is the center point detection using vanishing point. By prospective projection, all parallel lines in the object space except the lines parallel to the image plane will appear to intersect at a point called a vanishing point P as described in [32]. However this involves large scale computations and the goal is to have an on-chip mapping based on the scene of the image in the field of view of the eyelet with less computational overload.

The best method for an accurate on-chip sub-image center point computations is image centroiding. The image center is computed using the following equations [8]:

$$C_x = \sum \frac{I_i \times x_i}{I_i} \tag{2.31}$$

$$C_y = \sum \frac{I_i \times y_i}{I_i} \tag{2.32}$$

where C_x, C_y are the centroid points, I_i can be taken as the photo detector voltage at pixel i and x_i and y_i are horizontal and vertical position of the pixel i.

The possibility of on-chip image correction and reconstruction from the multiple sub-images would drastically reduce the bandwidth of the sensor, contributing to low power mode of operation. Such a system would enable us to reduce the volume of the imaging optics and pave the way to realize a compact image capturing system.

2.6 Summary

- Eyes can be broadly divided into simple aperture eyes and compound eyes. The single aperture eyes have single lens, while compound eyes have multiple lenses.
- Compound eyes are again classified into apposition eyes and superposition eyes. In apposition eyes each lens unit (ommatidium) acts as a separate eye, while in superposition light from different ommatidium are superimposed. Apposition eyes are well suited for diurnal animals while superposition eyes are most suited for nocturnal animals.
- The ommatidia are all arranged on a curved surface giving the compound eyes a wide field of view.
- The smaller lens diameter of the compound eyes prevent light from being focused and thus insect eyes suffer from diffraction limitation of the lens. The resolution obtained with compound eye is very poor compared to single aperture eye.
- The contrast can be increased either by increasing the lens aperture allowing more light to pass through or increasing the time for which the light is collected.
- The small lens aperture of the compound eyes gives them poor contrast vision. The response time of the animals is also very low thus cannot spend more time integrating light. These makes the overall resolution of the compound eye very poor.
- Light sensitivity can be increased by summing the electrical signal optically or neurologically, but this reduces both the spatial resolution as well as the temporal resolution.
- The temporal resolution, i.e. the number of images formed in a given time instance, is much higher for the insects eyes than the human eyes.
- The ommatidia provides a large focal depth because of extremely short focal length.
- In insect's eye the angular resolution is constant for a large distance, while the spatial resolution scales linearly.
- The insects mostly have trichromatic colour vision similar to humans. However their visible spectrum is shifted towards shorter wavelengths. Their eyes are more sensitive to ultraviolet light than to blue or green light.
- The flicker frequency of human eye is only 15-20 times a second while the flicker frequency of insects is very high, for example it is 300 for honeybee. Thus the temporal resolution of the insects eye is much higher than the human eye.
- These photoreceptors in the compound eyes are preferentially sensitive to a specific *e*-vector direction due to their strong dipole moment. This allows polarization vision in the compound eyes. Human eyes are polarization blind.
- In nature polarization vision is often used for tasks like contrast enhancement, camouflage breaking, object recognition, and signal detection and discrimination.
- The artificial compound eyes can be designed either using multiple camera systems where the scene is electronically distributed across multiple lenses or using single camera system where the scene is optically distributed across multiple lenses.
- Plenoptic function is used for the visual representation of the scene observed. A pinhole camera captures only one portion (pyramid) of the plenoptic function while a multichannel camera captures more than one portion of the plenoptic function.

- In artificial compound eyes the resolution and sensitivity are inversely related. The sensitivity increases with the diameter of the lens but the resolution decreases.
- For every field of view, diffraction is more important than geometrical aberrations.
- The micro-lens array can be fabricated using techniques like, ion exchange method, photosensitive method, thermal reflow method, chemical vapour deposition, hydrophobic method, focused ion beam method, hot embossing, micromachining etc. Thermal overflow method is most popular because of low cost and simple fabrication steps.
- An micro-lens array with a field of view of 124° and a maximum contrast of 0.3LP/° has been designed and demonstrated.
- The mapping of spherical angular distribution onto a square detector, causes overlapping of the field-of-view, this results in image distortions.
- Nonlinearities in the optical system can be classified as geometric distortion and photometric distortion. The geometric distortion is due to the variation in the focusing distance while photometric distortion is due to position dependent loss of light.
- A high resolution image can be obtained from the set of low resolution images by pixel rearranging.

References

[1] Nillison, D.E.: Vision optics and evolution. BioScience 39(5), 298–307 (1989)
[2] Land, M., Fernald, R.: The evolution of eyes. Annual Review of Neuroscience 15, 1–29 (1992)
[3] http://watchingtheworldwakeup.blogspot.com
[4] Land, M., Nilsson, D.: Animal Eyes. Oxford University Press, Oxford (2002)
[5] Land, M.F.: Visual acuity in insects. Annual Review of Entomology 42, 147–177 (1997)
[6] Srinivasan, M.V., Bernard, G.D.: The effect of motion on visual acuity of the compound eye: a theoretical analysis. Vision Research 15, 515–525 (1975)
[7] Land, M.F.: Optics of the eyes of phronima and other deep-sea amphipods. Journal of Computational Physiology 145, 209–226 (1981)
[8] Duparre, J.W.: Microoptical artificial compound eyes. PhD Dissertation (2005)
[9] Briscoe, A.D., Chittka, L.: The evolution of colour vision in insects. Annual Review in Entomology 46, 471–510 (2001)
[10] Nilsson, D.E., Labhart, T., Meyer, E.: Photoreceptor design and optical properties affecting polarization sensitivity in ants and crickets. Journal of Comparitive Physiology A-Sensory Neural and Behavioural Physiology 161, 645–658 (1987)
[11] Lambart, T., Meyer, E.: Neural mechanism in insect navigation: polarization compass and odometer. Current Opinion in Neurobiology 12(6), 707–714 (2002)
[12] Shashar, N., Cronin, T.W.: Polarization contrast vision in octopus. The Journal of Experimental Biology 199, 999–1004 (1996)
[13] Cronin, T.W., Shashar, N., Caldwell, R.L., Marshall, J., Cheroske, A.G., Chiou, T.H.: Polarization vision and its role in biological signaling. Intergrative and Comparative Biology 43(4), 549–558 (2003)

[14] Johnsen, S., Marshall, A.J., Widder, E.A.: Polarization sensitivity as a contrast enhancer in pelagic predators: lessons from in situ polarization imaging of transparent zooplankton. Philosophical Transactions of the Royal Society: B 366(1565), 655–670 (2011)
[15] Baker, P., Fermüller, C., Aloimonos, Y.: A spherical eye from multiple cameras (or how to make better models). In: Proceedings of Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 576–583 (2001)
[16] Richter, J.: The Notebooks of Leonardo da Vinci, vol. 1, p. 39. Dover, New York (1970)
[17] Adelson, E., Wang, J.: Single lens stereo with a plenoptic camera. IEEE Transaction on Pattern Analysis and Machine Intelligence 14(2), 99–106 (1992)
[18] Neumann, J., Fermueller, C., Aloimonos, Y.: Eyes from eyes: analysis of camera design using plenoptic video geometry. Tech. rep., Technical Report CAR-973 (2001)
[19] El-Sallam, A.A., Boussaid, F.: Spectral-based blind image restoration method for thin tombo imagers. Sensors 8, 6108–6124 (2008)
[20] Tanida, J., Yamada, K.: Thin observation module by bound optics (TOMBO): Concept and experimental verification. Applied Optics 40, 1806–1813 (2001)
[21] Fife, K., Gamal, A.E., Wong, H.: A 3d multi aperture image sensor architecture. In: Proceedings of Custom Integrated Circuits Conference, pp. 281–284 (2006)
[22] Duparre, J., Schreiber, P., Dannberg, P., Scharf, T., Pellia, P., Volkel, R., Herzig, H., Brauer, A.: Artificial compound eyes - different concepts and their application to ultra flat image acquisition sensors. In: Proceedings of MOEMS and Miniaturized Systems IV SPIE, vol. 5346, pp. 89–100 (2004)
[23] Duparre, J., Radtke, D., Tunnermann, A.: Spherical artificial compound eye captures real images. In: Proceedings of SPIE, vol. 6466(64660), pp. 1–9 (2007)
[24] Hessler, T., Rossi, M., Pedersen, J., Gale, M., Wegner, M., Tiziani, H.: Microlens arrays with spatial variation of the optical functions. Digest of Topical Meeting on Microlens Array 13, 42–47 (1997)
[25] Ko, H.C., Stoykovic, M., Song, J., Malyarchuk, V., Choi, W., Yu, C., Gegges III, J.B., Xiao, J., Wang, S., Huang, Y., Rogers, J.: A hemispherical electronic eye camera based on compressible silicon optoelectronics. Nature 454, 748–753 (2008)
[26] Duparre, J.W., Wippermann, F.C.: Micro-optical artificial compound eyes. Bioinspiration and Biommetics 1, 1–16 (2006)
[27] Moens, E., Meuret, Y., Ottevaere, H., Sarkar, M., Bello, D.S.S., Merken, P., Thienpont, H.: An insect eye-based image sensor with large field of view. In: Proceedings of SPIE, vol. 7716, pp. 7716–7785 (2010)
[28] Tardiff, J.P., Strum, P., Roy, S.: Plane based self calibration of radial distortion. In: Proceedings of International Conference on Computer Vision, pp. 1–8 (2007)
[29] Tudela, R., Brueckner, A., Duparre, J., Braeuer, A.: An image restoration approach for artificial compound eyes. In: Proceedings of Electronic Imaging SPIE, vol. 6812, pp. 68120O.1–68120O.1 (2008)
[30] Tsai, R.Y.: An efficient and accurate camera calibration technique for 3d machine vision. In: Proceedings of Conference on Computer Vision and Pattern Recognition, pp. 364–374 (1986)
[31] Tordoff, B., Murray, D.: Violating rotating camera geometry: The effect of radial distortion on self-calibration. In: Proceedings of International Conference on Pattern Recognition, vol. 1, pp. 423–427 (2000)
[32] Forsyth, D., Ponce, J.: Computer Vision a modern Approach, 2nd edn. Prentice Hall Inc., New Jersey (2003) ISBN: 9780130851987
[33] Liebe, C.C.: Solar compass chip. IEEE Sensors Journal 4(6), 779–786 (2004)

Chapter 3
Design of a CMOS Image Sensor

Image sensors have become significant due to the high demand from different applications. Charge-Coupled Device (CCD) and Complementary Metal-Oxide Semiconductor (CMOS) image sensors are two different technologies used for capturing images digitally. Both types of imagers are fabricated in silicon and convert light into electronic signals for processing.

Section 3.1 of this chapter discusses briefly the CCDs and CMOS image sensors. Section 3.2 gives an overview of the photodetection in silicon. The different types of photodiodes used to sense the light intensity are discussed. Section 3.3 discusses the basic types of pixels commonly used in CMOS image sensors. The advantages and disadvantages of each type are elaborated. Section 3.4 discusses the performance parameters of the CMOS image sensors and section 3.5 presents the peripheral circuits used in conventional CMOS image sensors. Section 3.6 describes the design of a 128x128 pixel CMOS image sensor chip with in-pixel analog memory and a 1-bit dynamic memory in a standard 0.18μm 1-poly-3-metal CMOS CIS (CMOS Image Sensor) process. The characterization and measurement results of the image sensor are also presented in section 3.7.

3.1 Introduction

The image sensors are electronic transducers which convert optical signals into electronic signals. This involves two steps: a) conversion of incident photon into electron-hole pairs and b) conversion of the electron-hole pair charge into voltage. The image sensors used in cameras are typically of two types, they are based on either a charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) technology. In CCDs and CMOS image sensors the process of the incident photon into electron-hole pair conversion is the same, however in CCD the charge is converted into voltage at the output during readout while in CMOS image sensors the charges are converted in the pixel itself as shown in figure 3.1. The vision or image sensor is similar to photographic film, which, when exposed to light, captures an image. It converts light into an electronic signal which is then processed and

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 49–104.
DOI: 10.1007/ 978-3-642-34901-0_3 © Springer-Verlag Berlin Heidelberg 2013

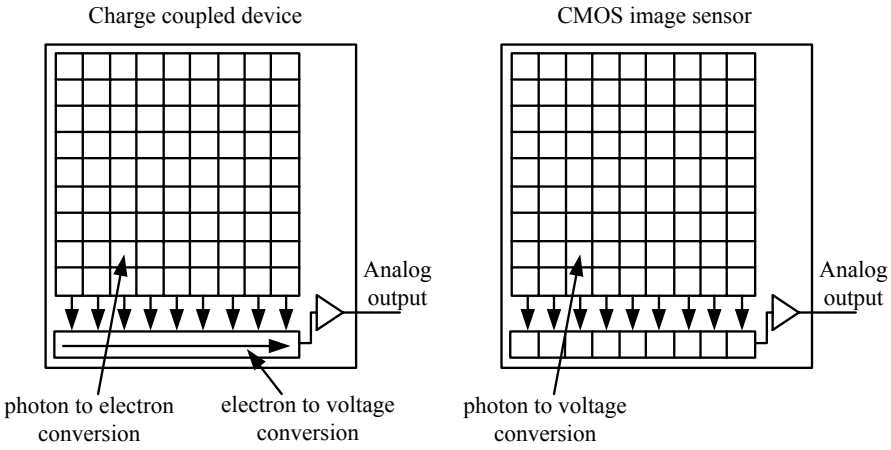


Fig. 3.1 Functional difference between a CCD and a CMOS image sensor

displayed. Both CCDs and CMOS sensors use arrays of millions of photo-detection sites, which are commonly referred to as pixels.

CCDs were invented in 1969 by George Smith and Willard Boyle at Bell Labs. These devices were originally conceived as a memory element for storing computer data and were not intended for capturing photographic images. A CCD is fabricated on a single crystal wafer of *p* or *n*-type silicon and consists of a one- or two-dimensional array of charge storage cells. The charge storage cells are simple series connection of metal-oxide-semiconductor capacitors that can be charged by light. It can hold a charge corresponding to variable shades of light, which makes it useful as an imaging device. Each storage cell has several closely spaced electrodes (gates) on top, separated from the silicon by an insulating layer of silicon dioxide. By applying a coordinated sequence of clock pulses to all the electrodes in the array, the charges accumulated are transported from one capacitor to the next until it reaches the output node where the charges are converted into voltage and measured.

CCDs however have several disadvantages. CCDs requires a manufacturing process which is different to that used for manufacturing memories and other processors. Thus for the fabrication of CCDs specialized fabrication units are required, making CCDs inherently more expensive. In CCDs the entire contents has to be read out and does not support region based readouts. This step by step process of readout due to limited output nodes, also reduces the speed at which the CCDs can operate. There are also a number of supporting chips required for the CCD sensor, each of which adds to the complexity and size of the camera design, increasing cost and power consumption.

CMOS or Complementary Metal Oxide Semiconductor was introduced as a vision sensor technology in the early 1990s. CMOS refers to a generic term for the process which is used to produce numerous other semiconductor items such as memories and processors. CMOS image sensors does not need a specialized fabrication unit and could be made with conventional semiconductor process.

With CMOS, a photodiode is typically used to generate charges, which are immediately converted into a voltage at the pixel. The supporting circuitry can thus be located alongside each photodiode or in the same chip, making the fabrication process cheaper compared to the CCDs. This reduces cost of supporting circuitry required, camera complexity, and also power consumption. The CMOS image sensors need lower voltage than the CCD sensors, and thus lower power consumption. The CMOS image sensors also allows for random access and thus region based readout is possible, allowing the CMOS image sensors to operate at very high speeds. The CMOS image sensors are easily capable of running at over 500 frames per second at megapixel resolution.

In the recent years, the CMOS image sensors have become dominant and are being used in many applications like, automobiles, digital camera, surveillance systems etc. CMOS image sensor have yet not been able to replace CCDs in the applications where very low noise operation is desired. The CMOS image sensors with their many amplifiers at each pixel, suffers from the fixed pattern noise, which arises due to the process variations while fabricating the amplifiers. CCDs are less susceptible to noise compared to CMOS and thus create higher quality, lower noise images. The CMOS image sensors are also less light sensitive compared to the CCDs. In CCDs, each individual units are only collecting light while the amplifiers inside the pixel in the CMOS image sensor reduces the light collecting area and thus are less sensitive to light. However the CMOS image sensors are being actively used as they are much less expensive and allow for the integration of processing circuitry with the photoreceptor, allowing for random accessibility, enhancing output bandwidth.

3.2 Photodetector

The photoelectric effect is the basic physical process of light detection in solid state photodetectors. The incident electromagnetic wave (light) is converted into an electrical signal by releasing and accelerating current-conducting carriers within the semiconductor [1]. In semiconductor such as silicon, there are few free electrons thus the valence band is completely filled and the conduction band is empty, and these two bands are separated by an energy gap. In order to generate an electrical signal, electrons have to be excited from the valence band into the conduction band as shown in figure 3.2.

When light is incident on a semiconductor, it may be reflected, absorbed or transmitted. The energy of the absorbed light particles (photons) forces the electrons to excite from the valence band into the conduction band when the absorbed energy is higher than the energy gap, thus producing electron-hole pairs. In silicon, the photon energy should be $\geq$1.124eV at 300°K, which corresponds to a light wavelength of λ=1050nm. The photons in the visible range have enough energy to generate electron-hole pairs in the silicon, the blue photon with a wavelength of 400nm have an energy of 3.1eV while the red photon with a wavelength of 700nm have an energy of 1.77eV. The number of electrons excited thermally from the valence band into the conduction band in silicon is given by:

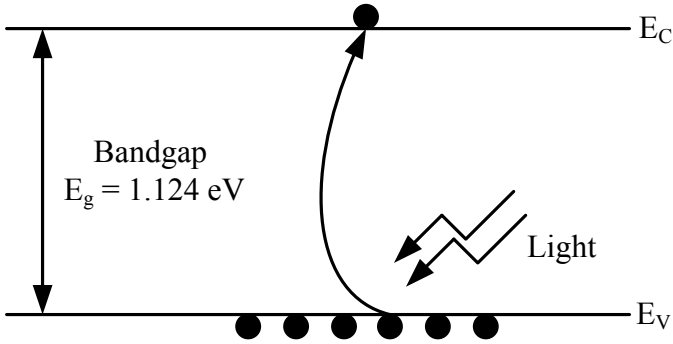


Fig. 3.2 Silicon energy bandgap

$$n_i = 2(\frac{2\pi m_e kT}{h^2})^{3/2} exp(\frac{E_g}{2kT})(\approx 10^{10}/cm^3 \ \ at \ \ T = 300^\circ K) \tag{3.1}$$

where n_i is the number of thermally excited electrons, *h* is the Plank's constant, *k* is Boltzmann's constant, *T* is the absolute temperature, E_g is the band gap energy and m_e is the effective mass of the electron.

Silicon based photodiodes are sensitive to light in the wide spectral range of 200-1200nm, which extends from deep ultraviolet through the visible to the near infrared. The most popular types of semiconductor based photodetectors used in image sensors are the reverse-biased *p*-*n* junction photodiode and the $p^+/n/p$ pinned photodiode. Though several other photosensitive structures are possible in a standard *n*-well CMOS process, here the discussion will be limited to *p*-*n* and $p^+/n/p$ junction photodiodes.

A *p*-*n* junction is formed when an *n*-type semiconductor is placed in contact with a *p*-type semiconductor. Since the *n*-type region has excess of electrons, the concentration gradient causes the electrons to diffuse from the *n*-type region to *p*-type region and the holes diffuse from the *p*-type to the *n*-type region. This diffusion process is continued till equilibrium is reached when the electric field of these carriers, is equal to the concentration gradient thus forming a junction. In order to function as a photodiode, the *p*-*n* junction is reverse biased. This forms a depletion region between the *p*-type region and the *n*-type region which is depleted of mobile charges. Light absorbed in the depletion region generates electron-hole pairs, which contribute to the generation of the photocurrent as shown in figure 3.3 and 3.4.

The electron-hole pair generation in silicon is also dependent on the penetration depth of the photons into the silicon. The amount of light absorbed in the silicon is given by its absorption coefficient α. In the visible spectrum, α is a decreasing function of wavelength. The intensity of light at the depth x traveled into the semiconductor can be expressed by [2]

$$I(x) = I_0 \exp(\alpha x) \tag{3.2}$$

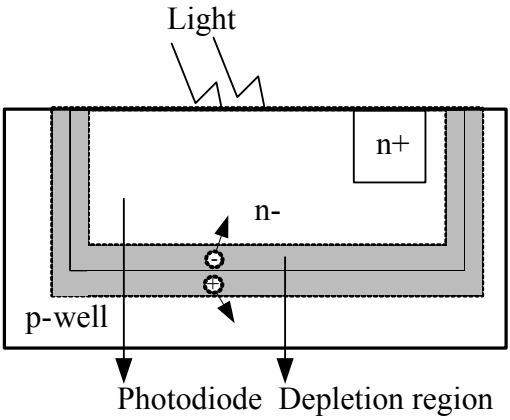


Fig. 3.3 n-well/p-well (sub) photodiode

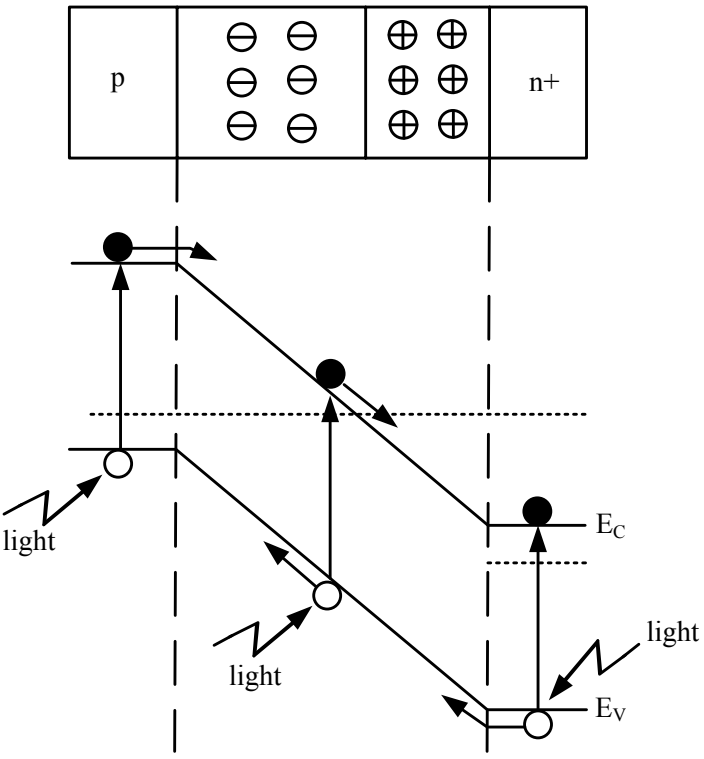


Fig. 3.4 Energy band diagram of a pn junction in reverse biased condition

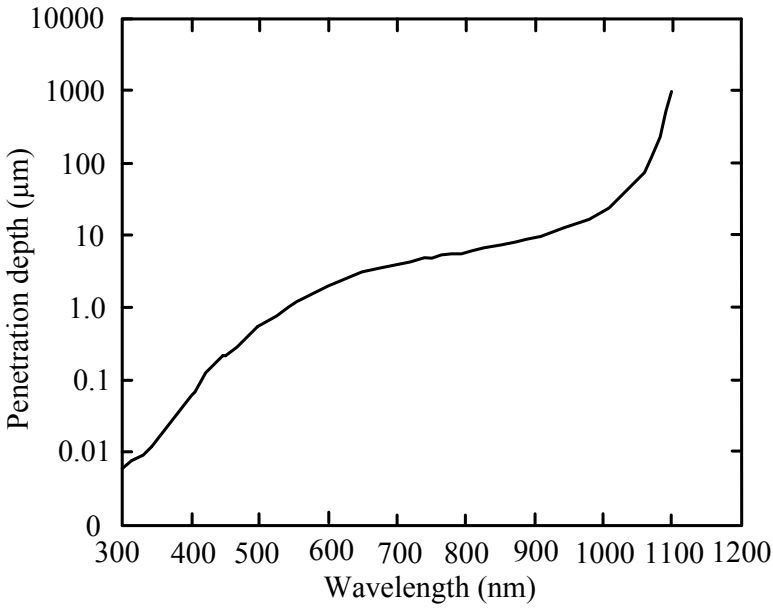


Fig. 3.5 Penetration depth versus wavelength in silicon

where I_0 is the light intensity at the surface. Figure 3.5 shows the absorption depth of a photon in silicon as a function of wavelength. It can seen from the figure that 99% of violet and blue light are absorbed within the penetration depth of approximately $x \approx 0.6\mu m$ and red light requires a depth of $x \approx 16.6\mu m$ for absorption.

The photodiode acts as a capacitor with the boundaries of the depletion region as the two plates of a parallel plate capacitor. At any applied reverse bias voltage$V$ the capacitance for a step junction is given by:

$$C = \frac{K_s \epsilon_0}{W_d} A_j = K_s \epsilon_0 A_j [2 K_s \epsilon_0 P \mu (V + V_{bi})]^{-1/2} \tag{3.3}$$

where W_d is the depletion region width, A_j is the diffused area, K_s the dielectric constant, ϵ_0 the permittivity of free space, P the resistivity of the substrate, μ the mobility of the majority carriers in the substrate, V_{bi} is the built-in voltage and V is the applied voltage across the junction.

Photodiodes are usually operated in photoconductive mode, where, a reverse voltage is applied to the diode and the resulting photocurrent is measured. The dependence of the photocurrent on the light power can be very linear over six or more orders of magnitude. The photocurrent of an ordinary photodiode is very small (in the range of femto- to picoamperes) under typical illumination conditions. Thus, a charge integration based operation is common in many active pixel sensors [3]. The basic idea of direct integration is illustrated in figure 3.6(a).

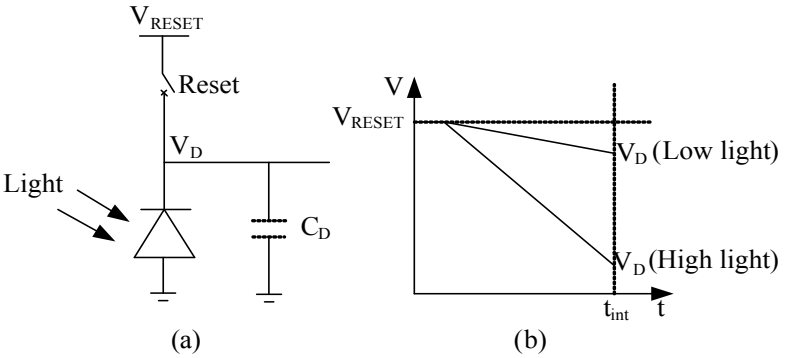


Fig. 3.6 (a) Pixel in integration mode of operation (b) charge versus time for two photocurrent value

The photodiode is first reset to a voltage V_{RESET} ($V_D = V_{RESET}$). The reset switch is then opened (disconnected) and the diode capacitance C_D is discharged by the photocurrent. This capacitance consists of the capacitance of any connected device at the node *V* plus the photodiode junction capacitance itself. We can write:

$$C(V)\frac{dV(t)}{dt} = -i_{photo} - i_{dark} \tag{3.4}$$

where i_{photo} is the photogenerated current and i_{dark} is the current generated in dark.

The photocurrent is directly proportional to the illumination levels, thus the resulting voltage drop across the photodiode capacitance C_D is proportional to the light intensity and the integration time. Thus to measure the light intensity, the change in the voltage across the photodiode after a certain integration time (t_{int}) is measured. Examples of the voltage versus time characteristics for low light and high light conditions are shown in the figure 3.6(b) [4].

The pinned photodiode $p^+/n/p$ provides an improved photoresponsivity compared to standard *p-n* photodiode [5] due to its low dark current and good blue response. The concept of pinned photodiodes is borrowed from surface pinned interline CCD structures. It needs additional wafer implantation steps and an additional transistor (transfer gate) compared to the standard *p-n* photodiodes to improve noise performance, particularly reset noise, to reduce image lag and to suppress the dark currents [6]. The additional transistor allows extended functionalities such as electronic shuttering and dynamic range extension [7]. The cross section of a pinned photodiode is shown in the figure 3.7, showing also the transfer gate *(TG)*.

In a pinned photodiode the n^- terminal is separated from the surface by a p^+ pinning layer. The p^+ pinning layer is pinned to the substrate voltage. The *p^+/n-well* photodiodes where the p^+ in enclosed within its own *n-well* is not used as pinned photodiode. The $p^+/n-well$ photodiodes are not suitable for dense arrays because of the huge separation required between two wells. The passivation of defect and

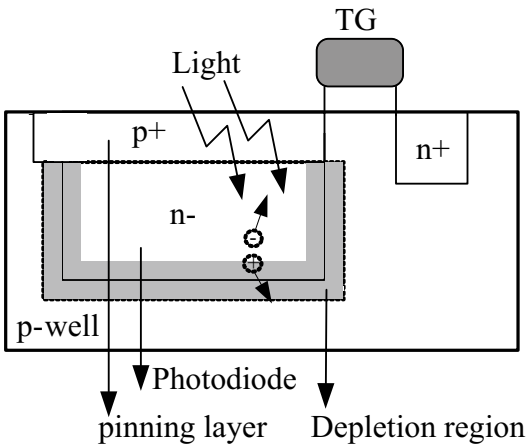


Fig. 3.7 p^+/n-well/p-well(sub) photodiode

surface states at the Si/SiO_2 interface in the pinned photodiodes contributes to the lowering of the thermal noise and dark currents as explained later. The transfer gate separates the floating diffusion node from the photodiode. This allows for the optimization of the floating diffusion node capacitance by sizing the *TG* transistor. The transfer gate is an asymmetric device which couples the photodiode and the floating diffusion node.

The pinning voltage for the p^+ pinning layer is 1V for the selected technology. When the photodiode is reset to the pinning voltage, a full depletion of the photodiode is achieved. In the full depletion conditions the depletion regions of $p^+ - n^-$ and $n^- - p$ meet, eliminating the majority carrier in the *n* region. A full depletion condition of the pinned photodiodes is preferred, to suppress dark currents and to reduce the *kTC* noise component due to the mobile electrons left after reset. If the pinned photodiode is not fully depleted, the mobile electrons left would contribute to additional *kTC* noise after reset. However since the pinning voltage of the photodiode is fixed, the charge transfer efficiency from the photodiode to the floating diffusion node is low which lowers the saturation charge or well capacity of the pinned photodiodes. Well capacity is defined as the maximum amount of charge (in electrons) that can be held by the integration capacitance of the photodiode. The well capacity of a pinned photodiode can be represented by $Q_{PD}=(V_{pinning} - V_{blooming})C_{PD}$, where $V_{pinning}$ is the pinning voltage, $V_{blooming}$ is the blooming voltage of the pixel and C_{PD} is the capacitance of the photodiode. The reduction of $V_{blooming}$ is limited by the minimum voltage required for stopping the photo-generated electrons from overflowing into the silicon substrate during high illuminations. Therefore, to increase the well capacity of the pixel, $V_{pinning}$ or C_{PD} should be increased. However, it is very difficult to increase the charge storage capacitance as the pixel size shrinks and a fixed $V_{pinning}$ results in a low well capacity. At the end of the exposure time, photogenerated carriers are accumulated and transferred to the floating diffusion $n^+(FD)$ by opening the transfer gate *TG*.

3.3 Basic CMOS Image Sensor Architecture

The architecture of a CMOS image sensor is shown in figure 3.8.

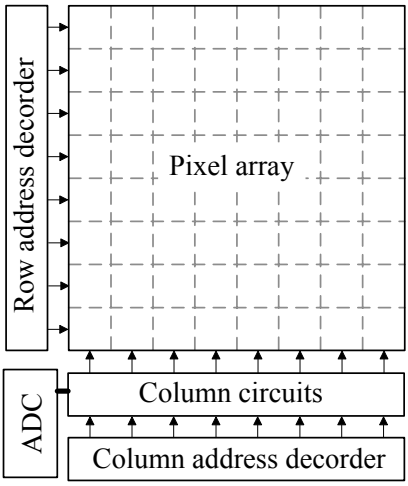


Fig. 3.8 CMOS image sensor architecture

It contains a 2-dimensional array of pixels and peripheral circuits. The CMOS image sensor image quality depends on the type of the pixel used in the imaging array. The typical pixel types are

- Passive pixel sensor (PPS)
- Active pixel sensor (APS)
- Digital pixel sensor (DPS)

3.3.1 Passive Pixel Sensor

A passive pixel has a photodiode and a MOS transistor which connects the photodiode output to the data line used for reading the pixel as shown in figure 3.9. When the light falls on the photodiode the charges are integrated in the photodiode capacitance, and when the MOS transistor is switched on, the accumulated charges are transferred to the column bus. The charges in the column bus are then converted into voltage using a charge amplifier.

For the case of n^+*/p-well* photodiode, the pixel can be made very compact if the n^+ of the MOS transistor is extended to serve as a photodiode. The main advantages of passive pixels are the high fill factor, large voltage swing and layout simplicity. The fill factor is defined as

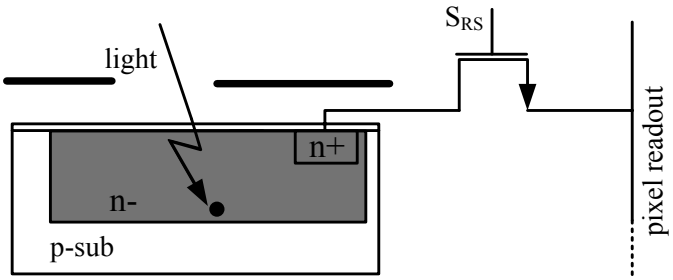


Fig. 3.9 Passive pixel

$$Fillfactor = \frac{\text{Area of the photodiode}}{\text{Area of the total pixel}} \tag{3.5}$$

The presence of only one transistor allows most of the pixel area to be active area collecting the photons. The main disadvantage of the passive pixel is that the photodiode junction capacitance adds to the parasitic column bus capacitance which results in longer response time and excess noise charge. The noise charge in electrons is expressed as

$$q_{noise} = \frac{\sqrt{kTC}}{q} \tag{3.6}$$

where k is the Boltzmann's constant (1.38 x 10^{-23} J/K), T is the absolute temperature, C is the capacitance and q is the elementary charge (1.6 x 10^{-19})C.

3.3.2 Active Pixel Sensor

In active pixel sensor, the column amplifier of the PPS is taken inside the pixel. The APS consists of a photodiode, reset transistor, in-pixel amplifier and a row select transistor as shown in the figure 3.10. This pixel is also often referred to as 3-Transistor pixel as the pixel consists of three MOS transistors. The photodiode can either be a *n+/p-well* or *n-well/p-well* junction. APS is the most commonly used pixel structure.

The pixel operation begins with a reset of the photodiode to V_{DD} through the reset transistor S_{RST}. After the reset, when light falls on the photodiode, the photo-generated charges discharge the photodiode capacitance causing the photodiode node potential to decrease. Thus the photo-generated charges are converted to voltage in the pixel itself. The conversion gain, which is defined as the output voltage per signal electron detected and expressed as $\mu V/e^-$, is determined by the photodiode capacitance. The in-pixel amplifier is often a source follower which amplifies the input signal. The gain of the source follower is given by

$$A_{vsf} = \frac{1}{1 + \frac{g_{mb}}{g_m}} \tag{3.7}$$

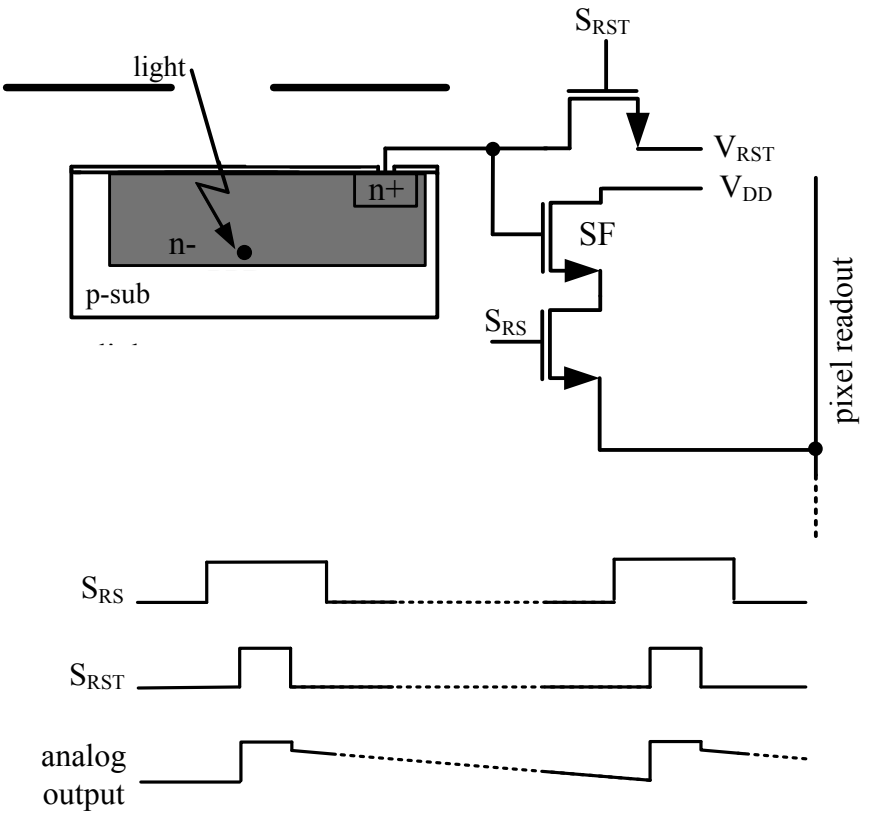


Fig. 3.10 3-Transistor active pixel

where A_{vsf} is the gain of the source follower, g_m is the transconductance and g_{mb} is the bulk transconductance of the source follower. After the signal has been amplified, it is readout through the row select (S_{RS}) switch. Since three transistors are used in APS, the fill factor is lower compared to PPS. However, because the photodiode can be reversely biased using a strong positive potential, a larger depletion region is formed which improves the quantum efficiency and the well capacity of the photodiode.

The disadvantage of the APS is limited output voltage swing. The highest output voltage is limited by two threshold voltage drops due to the NMOS reset transistor and the source follower. The lower output voltage is limited by the condition that the column load transistor needs to be biased in saturation. The minimum and maximum output voltage of the APS is given by equations (3.8) and (3.9)

$$V_{out,min} = V_{DS,sat} = V_{GS,M4} - V_{TH} \tag{3.8}$$

$$V_{out,max} = V_{DD} - V_{TH} - V_{GS,sf} \tag{3.9}$$

The mismatches in the amplifier in the APS results in fixed pattern noise *(FPN)*. The fixed pattern noise is the spatial variation in the signal from pixel to pixel due to process variations and is described in section 2.4. The APS also have a high temporal noise.

The limited output voltage swing of the APS is solved using a logarithmic pixel as shown in figure 3.11. Pixels in logarithmic mode operate continuously and convert the logarithm of the photocurrent into a corresponding voltage without the integration process. In the logarithmic pixel, the reset transistor in the figure 3.10 is diode connected. The very small photocurrent through the diode biases the reset transistor in subthreshold regime. In the subthreshold regime, the MOSFET drain current I_d is very small and exponentially increases with the gate voltage V_G

$$I_d = I_0 \exp\left(\frac{q}{nkT}(V_G - V_{TH})\right) \tag{3.10}$$

where I_0 is the current flowing when $V_G = V_{TH}$ and is a process dependent parameter. It is a function of the gate oxide capacitance and the aspect ratio of the reset transistor, q is the elementary charge, k is the Boltzmann's constant, T is the absolute temperature, n is the subthreshold slope, V_G is the applied gate voltage and V_{TH} is the threshold voltage. From equation (3.10), the photodiode potential (source potential of the transistor) can be written as

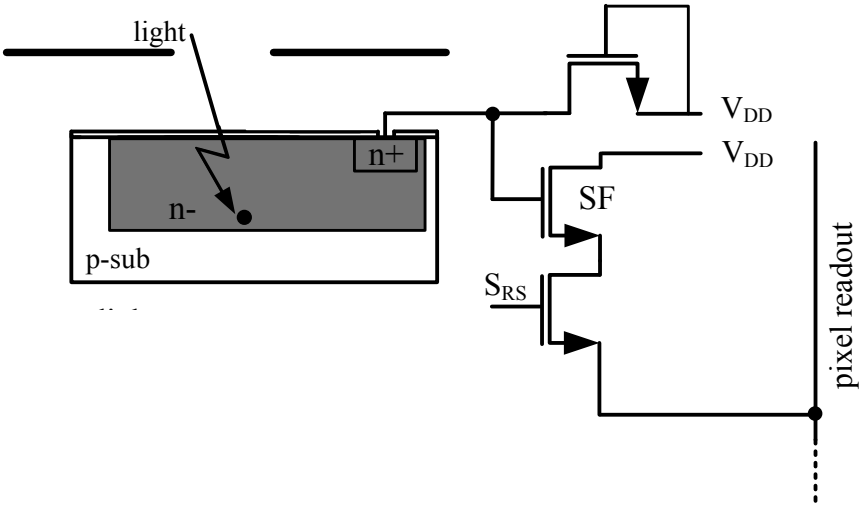


Fig. 3.11 Logarithmic pixel

$$V_{PD} = (V_{DD} - V_{TH}) - \frac{nKT}{q} \ln\left(\frac{I_p}{I_0}\right) \tag{3.11}$$

where I_p is the photodiode current which is equal to the drain current of the MOSFET. In the logarithmic pixel, the sense node varies logarithmically with the photocurrent, which provides pixel with a very high dynamic range but restricted voltage swing. However the photosensitivity in low light is very low compared with the 3-Transistor active pixel. Further since the pixel is not reset, the pixel suffers from image lag and is very noisy. The logarithmic pixels are also not suitable for high speed imaging, as the response time is very low specially in low light environment.

A 4-Transistor pixel has four transistor inside the pixel and is schematically shown in figure 3.12. In a standard 4-Transistor pixel, the pinned photodiode *(PPD)* collects the charges generated by the photons and a transistor with a transfer gate *(TG)* is used for transferring the collected charges from the photodiode to the floating diffusion *(FD)* node. The pixel also includes a transistor (S_{RST}) for resetting the *FD* to a predetermined voltage level (V_{RST}) prior to the charge transfer. The other two transistors are the source follower and row select transistor (S_{RS}). The row select transistor loads the column bus with the pixel output for readout.

The pinned photodiode forms a fully depleted *n*-region, which is "pinned to a voltage", V_{pinned}. This is the potential when the depletion region of the two junctions p^+–n^- and n^-–*p–sub* merge, thus depleting the entire n^- region. The fully depleted structure helps in the reduction of the reset noise and the problem of image lag. Reset noise is a thermal noise which occurs when the photodiode is reset to a different level each time the photodiode is reset. The image lag occurs when not all collected electrons are transferred to the FD node. Since the n^- gets fully depleted, the collected electrons from the previous frame are all lost and thus reduces the image lag.

In this region the photo-generated charges are collected and stored till the readout starts by transferring the charges to the *FD*. In a 4-Transistor pixel configuration, the potential well in the *PPD* forms the storage well, while a low voltage on the *TG* forms the barrier region preventing the electrons from exiting the *PPD* region.

In the pinned photodiode, the surface is inverted using the p^+ pinning layer. The inversion of the surface helps to reduce the electrons in the interface traps at the surface and thus the leakage current from the thermal generation of the interface states is reduced. The heavily doped p^+ layer contributes to the sensitivity of the photodiode to blue light, which generates electrons near the surface of the photodiode.

Since the photodiodes are pinned to a pinning voltage, the well capacity is reduced. The well capacity defines the maximum number of electrons that the potential well of the photodiode can store.

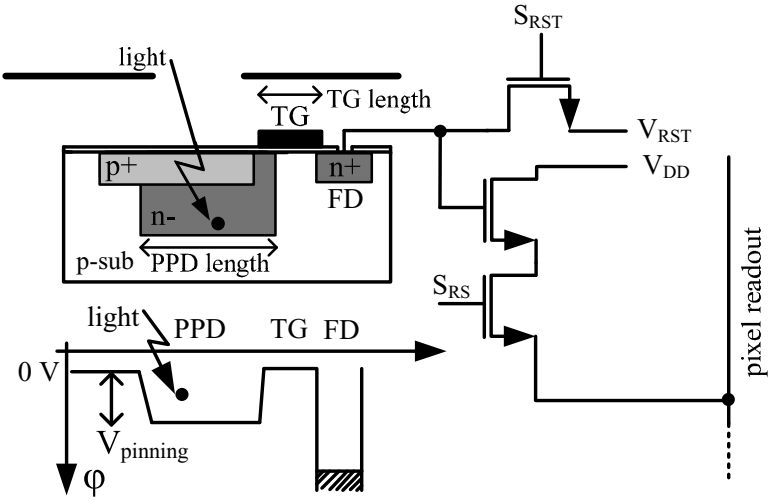


Fig. 3.12 4-Transistor pixel

3.3.3 Digital Pixel Sensor

The digital pixel sensor (DPS) integrates the *ADC* at the pixel level. In the *APS*, the *ADC* is performed using a per-array or per-column readout scheme. The DPS allows massive parallel analog to digital conversion at the pixel level allowing for pixel level image processing. The pixel level image processing is very useful in applications like automotive and surveillance systems.

Due to the pixel analog to digital conversion, DPS can offer improved noise figure and dynamic range. Two DPS implementation based on pulse width modulation (PWM) and pulse frequency modulations (PFM) to digitize the photodiode voltage are shown in the figure 3.13 and 3.14 respectively.

In PWM DPS, the photodiode photocurrent is converted into time measurement. The PWM scheme uses a comparator to compare the pixel output signal with the reference signal. PWM produces an output signal when photodiode output signal reaches a threshold value defined by the reference voltage to the comparator. When the comparator output is produced, the pixel is reset and the accumulation process starts again. The frequency (width) of the output signal is proportional to the input light intensity. Simple inverters can also be used in place of the comparator to minimize the area of a pixel. The input reference voltage to the comparator can either be a constant DC voltage or a ramp waveform. PFM as shown in figure 3.14 consists of series of inverters in the pixel. When the photodiode potential is lower than the threshold voltage of the inverter, the inverter chain in turned on and output pulse is produced. When the output pulse is produced, the photodiode is reset again and the accumulation process continues. The frequency of the inverter output is proportional to the input light intensity. *PFM* based photosensors are more frequently used in biomedical application, specially in low light environment.

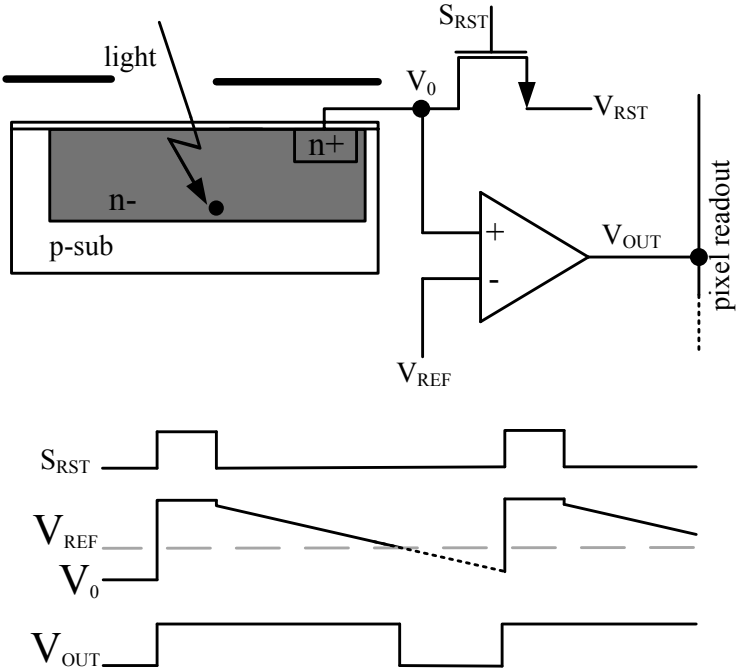


Fig. 3.13 DPS in pulse width modulation (PWM)

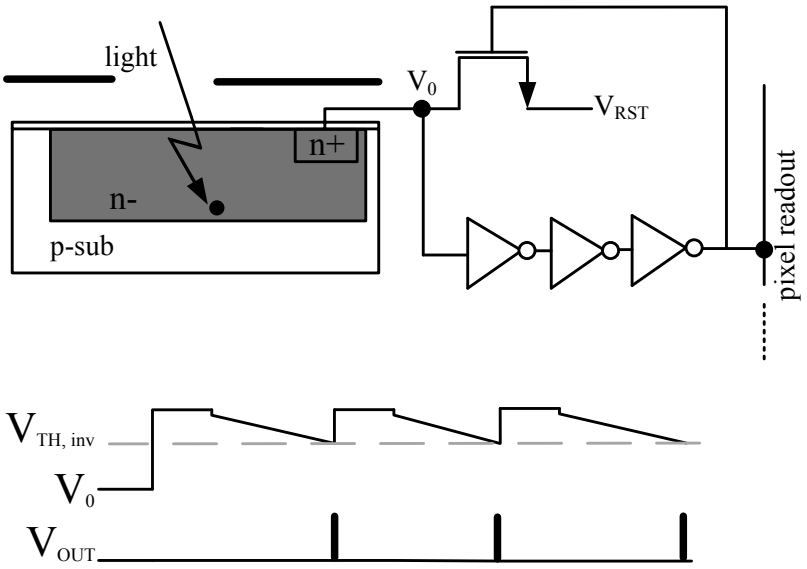


Fig. 3.14 DPS in pulse width modulation (PFM)

However the PWM and PFM based image sensor comes with a disadvantage of requiring more number of transistors per pixel, thus are difficult to implement for high resolution imaging. The higher number of transistors per pixel reduces the fill factor of the pixel, reducing the light sensitivity of the pixel. Both PFM and PWM are not suitable for low power applications, as the switching activity at the pixel level consumes a lot of power.

3.4 CMOS Image Sensor Performance

The performance of a CMOS image sensor are defined by numerous parameters. This section discusses few of the important performance parameters of CMOS image sensors such as quantum efficiency and spectral sensitivity, signal-to-noise ratio, dynamic range, pixel conversion gain, dark current in the pixel and the noises affecting the image sensor.

3.4.1 Quantum Efficiency and Spectral Responsivity

The quantum efficiency represents the imager's ability to intercept incident photons and to generate and collect signal electrons generated through the photoelectric effect. The quantum efficiency for visible wavelength is defined as the number of signal electrons generated for a given number of incident photons. The quantum efficiency is given as

$$QE = \frac{N_{sig}}{N_{photon}} \tag{3.12}$$

where N_{sig} is the number of electrons generated by the incident photons and N_{photon} is the number of photon incident on the light detection surface. For monochrome light the number of incident photons is given by

$$N_{photon} = \frac{\lambda}{hc} P A_{pix} t_{int} \tag{3.13}$$

where λ is the wavelength of incident light, h is the Planck's constant (6.62×10^{-34}) Js, c is the speed of light, P is the face-plate irradiance, A_{pix} is the pixel area and t_{int} is the integration time over which the photons are intergraded in the photodiode.

Normally spectral responsivity is used to characterize the light sensitivity of an image sensor. Spectral responsivity can be expressed as

$$R = \frac{I_{ph}}{P_{ph}} = \frac{q\lambda}{hc} QE \tag{3.14}$$

where R is the spectral responsivity I_{ph} is the photocurrent generated, P_{ph} is the incident optical power and QE is the quantum efficiency. The spectral responsivity is dependent on the wavelength of the incident light as seen from the equation (3.14). The spectral response of an conventional CMOS image sensor is shown in figure 3.15. The photoelectric effect is dependent upon a critical wavelength over which the photons do not have sufficient energy to excite the electrons from the valance band to the conduction band. When the bandgap energy of silicon becomes greater than the energy of the photon, the photons pass the silicon substrate without causing the photoelectric effect. Thus silicon behaves as transparent material for higher wavelengths with lower photon energy. The spectral responsivity thus considerably degrades for longer wavelength of light.

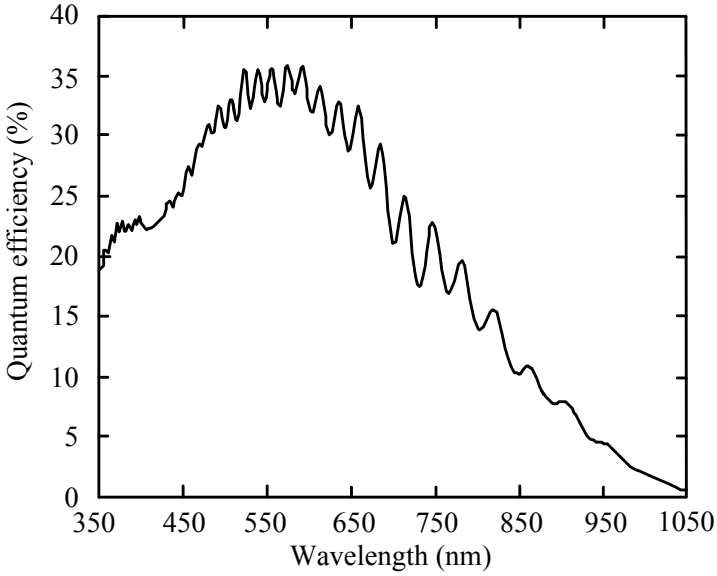


Fig. 3.15 Sample spectral response of a CMOS image sensor

There are several ways of improving the quantum efficiency and spectral responsivity. The fill factor which is the ratio of the photodiode to the pixel area should be as high as possible to be able to collect more number of available photons. A micro-lens is often used to improve the photon collection and thus the quantum efficiency. A micro-lens focuses the available light onto the optically active area of the pixel and thus improves in collecting the available photons. The quantum efficiency is also lowered by the absorption and reflection of light from the overlayering structure over the photodiode like the metal interconnects and oxide layers. To compensate for the photon loss due to absorbtion or reflection from the overlayers, backside illumination is used. In backside illumination the light is incident from the back of the silicon and thus a 100% quantum efficiency can be achieved. In addition, applying anti-reflection coating layers on top of the sensor can

minimize reflection loss. It is however difficult to produce coatings that are effective across the entire visible range. The coating that increases the spectral response in the longer wavelength often decreases the absorption of the shorter wavelengths.

Further the penetration depth of the incident photons depends on the wavelength of the light. The longer wavelength light can penetrate deeper compared to shorter wavelength. Thus the quantum efficiency depends on the width of the depletion region formed. Most of the photons with a wavelength between 450 and 700nm are absorbed either in the depletion region or within the silicon bulk depending on the depletion width. The photons absorbed in the depletion region exhibits higher quantum efficiency, as the electron-hole pairs generated within the depletion region move by drift process and are efficiently collected by the collection node. The photons absorbed in the bulk produces electron-hole pair which either recombine with holes or diffuse in the depletion region based on the diffusion length and carrier life time of the carriers in silicon. The electrons with a large diffusion lengths are able to diffuse into the depletion width are able to be collected by the collection node and thus increases the quantum efficiency.

3.4.2 Dynamic Range and Signal-to-Noise Ration (SNR)

The dynamic range of a sensor expresses the ability to capture dark as well as bright sections within an image while maintaining the highest fidelity. Technically expressed, the dynamic range is the ratio of the largest possible signal divided by the smallest possible signal that is still detectable. A low dynamic range would mean a loss of perceived intensity information. An increased dynamic range can be achieved by either lowering the noise floor or increasing the full well capacity as shown in the figure 3.16.

For a linear image sensor, the dynamic range is the ratio of the largest possible signal divided by the smallest possible signal it can detect. The largest possible signal is directly proportional to the full well capacity of the pixel. The lowest signal is the noise level when the sensor is not exposed to any light, also called the “noise floor”. Dynamic range *(DR)* can be expressed as [8]

$$\begin{aligned} DR &= 20\log_{10}\frac{i_{max}}{i_{min}} \\ &= 20\log_{10}\frac{Q_{sat}-(i_{dc}\times t_{int})}{\sqrt{(q\times i_{dc}\times t_{int})+\sigma^2_{read}+\sigma^2_{reset}+\sigma^2_{FPN}}} \end{aligned} \tag{3.15}$$

where Q_{sat} is the well capacity, i_{dc} is the dark current generated by the photodiode, t_{int} is the integration time, q is the elementary charge, σ^2_{read}, σ^2_{reset} and σ^2_{FPN} are the variances of the readout noise, reset noise and fixed pattern noise *(FPN)* respectively.

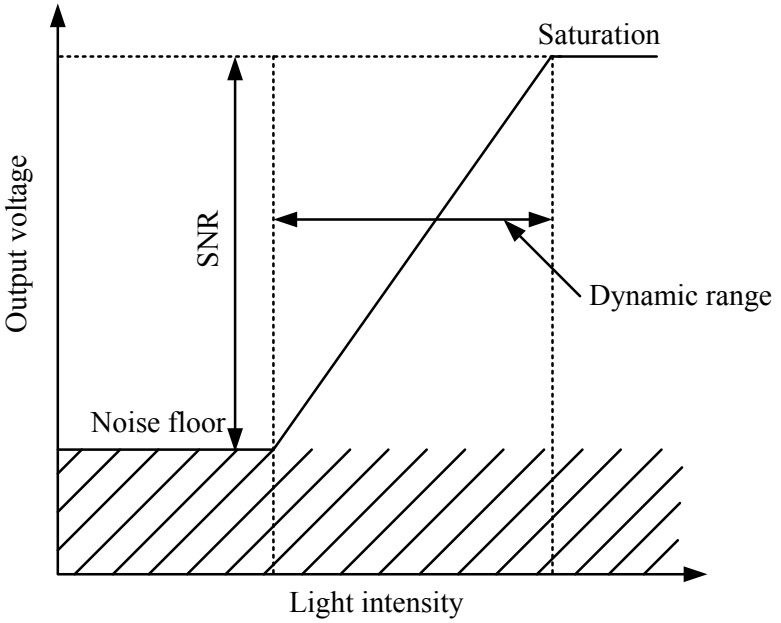


Fig. 3.16 Dynamic range and SNR of a CMOS image sensor

Signal-to-Noise ratio (SNR) as shown in the figure 3.16 is the ratio of the signal and the noise that can be measured for a given input light level. The SNR is expressed as

$$SNR = 20\log_{10}\left(\frac{N_{sig}}{n_{noise}}\right)[dB] \tag{3.16}$$

where N_{sig} is the number of electrons produced by the incident signal and n_{noise} is the total noise generated. When the photon shot noise is dominant then the SNR can be written as

$$SNR = 20\log_{10}\left(\frac{N_{sig}}{n_{photon}}\right)[dB] \tag{3.17}$$

The maximum photon shot noise is the square root of the signal level itself ($\sqrt{N_{sig}}$), and the maximum signal is the saturation signal N_{sat}. Thus from equation (3.17) the maximum SNR can be derived as

$$\begin{aligned} SNR &= 20\log_{10}\left(\frac{N_{sat}}{\sqrt{N_{sat}}}\right)[dB] \\ &= 10\log_{10} N_{sat}[dB] \end{aligned} \tag{3.18}$$

A high signal-to-noise ratio is important in order to obtain images of uniform segments and in order to individually perceive them as a realistic scene. Assuming photon-shot noise limited signal detection the maximum achievable signal-to-noise ratio benefits from the full well capacity in the sense that it increases proportionally to the square root of the signal itself. The SNR is a fundamental criterion for image quality in terms of noise. The dominant source of noise changes with change in illumination levels, thus SNR is dependent on the illumination conditions.

3.4.3 Pixel Conversion Gain

The output of the image sensor is measured in voltage at the output node. However, the signal generated in the photodiode is in the form of number of electron-hole pairs generated due to the incident light. Thus there is a charge-to-voltage conversion in the pixel. The pixel conversion gain quantifies the conversion process of the generated charges into voltage. The conversion gain of a pixel is defined as the output voltage produced for one electron detected at the charge detection node (floating node). It is usually expressed as $\mu V/e^-$. The conversion gain is expressed as

$$CG = \frac{q}{C_{FD}} \tag{3.19}$$

where q is the elementary charge and C_{FD} is the capacitance at the charge detection node. Almost all the parameters of the CMOS image sensor like response linearity, light sensitivity, uniformity and noise of the pixels are all influenced by the conversion gain. Thus making the conversion gain to be one of the most important parameters of a CMOS image sensor.

3.4.4 Dark Current

Dark current is the current flowing through the photodiode even in the absence of light signal. The dark current is a leakage current and is required to be minimized for optimal performance of the pixel. In a $n-well/p-sub$ photodiode, there are three sources of dark current

- thermal generation of electrons in the depletion region
- thermal generation of electrons in the bulk region
- generation of electrons at the interface states of the surface

The total dark current density in a CMOS image sensor is a combined contribution from the thermal generation in the depletion region, bulk region and surface of the silicon. The total dark current density is expressed as

$$J_{dark} = J_{depletion} + J_{diffusion} + J_{surface} \tag{3.20}$$

where J_{dark} is the total dark current density of the pixel, $J_{depletion}$ is the dark current contribution from the depletion region, $J_{diffusion}$ is the dark current contribution from the bulk region and $J_{surface}$ is the dark current contribution from the surface states. Thermal generation describes how the free minority carriers are generated. In CMOS image sensor, the photodiode is usually reverse biased to have a wide depletion region. The wide depletion region helps in collecting long wavelength light and thus improves the quantum efficiency. The electric field across the depletion region is usually high and thus any minority carriers generated within the depletion region, are efficiently collected through a drift process. In thermal generation, the minority carriers are thermally generated from the generation recombination process in the silicon. The electron transition from the valance band to the conduction band is often not direct rather indirect through the intermediate energy state (defects) located in the silicon bandgap [9]. These intermediate energy states assist the electrons in the valance band being excited to the conduction band. These electrons in the conduction band constitutes the dark current.

The dark current caused by the thermal generation in the depletion region is:

$$J_{depletion} = \int_0^W qGdx \approx qGW = \frac{qn_iW}{\tau_g} \tag{3.21}$$

where $J_{depletion}$ is the dark current density in the depletion region, W is the depletion width, q is the electronic charge, G is the generation rate, n_i is the intrinsic concentration, and τ_g is the generation lifetime and is given by

$$\tau_g = \frac{2cosh\left(\frac{E_t - E_i}{kT}\right)}{v_{th}\sigma_o N_t} \tag{3.22}$$

where E_t is the defect energy level, E_i is the intrinsic energy level, k is the Boltzmann's constant, T is the absolute temperature, v_{th} is the thermal velocity σ_o is the capture cross section and N_t is the density of the generation centers.

Since the properties of silicon defects are determined by the fabrication process, the only design parameter which can control the dark current density is the depletion width as shown in equation (3.21). The depletion width can be controlled by the amount of doping concentration and the reverse bias voltage applied. To have a lower dark current density, a smaller depletion width is desired, however a smaller depletion width lowers the quantum efficiency as well. The most efficient way to improve the performance is to optimize the technology and reduce the defect densities.

The minority carriers can also be generated at the neutral region (bulk region) and diffuse into the depletion region to contribute to the total dark current. The minority carriers generated in the depletion region follow drift process while the minority carriers generated in the bulk follow diffusion process. The diffusion process is less efficient compared to drift process and depends strongly on the doping concentration. A higher doping concentration lowers the time constant and

diffusion length of the minority carriers and thus the probability of recombination increases. The dark current density in the bulk due to diffusion can obtained using the equation

$$J_{diffusion} = \frac{qD_n n_{p0}}{L_n} = q\sqrt{\frac{D_n}{\tau_n}}\frac{{n_i}^2}{N_A} \tag{3.23}$$

where D_n is the electron diffusion coefficient, n_{p0} is the electron concentration in the *p*-type region, L_n is the electron diffusion length, τ_n is the electron lifetime, n_i^2 is the intrinsic carrier concentration and N_A is the acceptor doping concentration. The diffusion current as seen from equation (3.23) is proportional to the square of the intrinsic concentration n_i. The minority carrier generation in the bulk region is due to the presence of silicon defects in the bulk region. However in the absence of an electric field, the generation-recombination process generating the dark current is less significant.

The dark current generation at the surface is also due to thermal generation. At the surface due to the oxide stress and the presence of dangling bonds (energy states) of the oxide layer, the density of surface defects is much larger compared to the defect density in the depletion and bulk region. The density of the surface states is determined mainly by the fabrication process. In modern CMOS image sensor, the surface dark current generation dominates over other dark current generating mechanisms. The dark current density due to surface generation can be expressed as

$$J_{surface} = \frac{qS_0 n_i}{2} \tag{3.24}$$

where S_0 is the surface generation velocity, which is determined by the trap energy and capture cross section.

3.4.5 Noise in CMOS Image Sensor

Noise determines the minimum amount of light that can be detected. The noise sources in the CMOS image sensor can be characterized into either random noise or pattern noise. The random noise is temporally random and is not constant from frame to frame, thus can be removed by averaging. Pattern noise is the variation in the signal from pixel to pixel under uniform illumination. The pattern noise does not change from frame to frame, thus cannot be reduced by averaging.

3.4.5.1 Temporal Noise

The temporal noise sets the fundamental limit on image sensor performance, especially under low illumination condition. The temporal noise can be categorized into a number of noise components like

Shot Noise - The shot noise is a white noise which arises from the discrete nature of the electrons themselves. The random generation of carriers can either be due

to thermal generation (shot noise of the dark current), and is often called as dark current shot noise or due to the random generation of the photo electrons caused by the random arrival of of the photons itself, often referred to as photon shot noise. The photo generated charges for a uniform light source will have a random Poisson distribution [10]. The Poisson distribution is represented by

$$P_N = \frac{(N_{avg})^2 - e^{-(N_{avg})}}{N!} \tag{3.25}$$

where N and N_{avg} are the number of particles and the average, respectively. The Poisson distribution has an interesting property that the variance is equal to the average value, or

$$< N_{shot}^2 >= (N - N_{avg})^2 = N_{avg} \tag{3.26}$$

The power spectral densities of the shot noise is constant over all frequencies and hence referred to as "white noise". The dark current shot noise is noise read-out when the device is not exposed to light. It is due to the random fluctuation of the dark current value. The dark current shot noise arises due to the presence of intermediate energy states between the silicon bandgap. The intermediate energy states are introduced by the presence of defects. The defects trap and detrap electrons, which introduces a randomness in the signal measured. The dark current shot noise depends on the technology used, the temperature, the doping concentration and the reverse bias voltage applied through the depletion layer width. The dark current shot noise is given by the equation

$$< n_{dark}^2 >= \frac{J_{dark} A t_{int}}{q} \tag{3.27}$$

where n_{dark}^2 is the dark shot noise variance, J_{dark} is the dark current density, A is area of the photodiode, t_{int} is the integration time and q is the elementary charge.

The photon shot noise is given by the equation

$$< n_{photon}^2 >= \eta I_0 A t_{int} \tag{3.28}$$

where n_{photon}^2 is the photon shot noise variance, η is the quantum efficiency, I_0 is the incident photon flux, A is the area of the photodiode and t_{int} is the integration time.

The total shot noise is given by adding equations (3.29) and (3.28)

$$< n_{shot}^2 >=< n_{dark}^2 > + < n_{photon}^2 > \tag{3.29}$$

Reset Noise - The reset noise originates from the thermal noise of the reset switch in the pixel. When the reset transistor is turned ON, it can be considered as a resistor which contain thermal noise. The reset voltage is sampled on the photodiode capacitance when the reset transistor is ON. The sampling operation will also sample noise, which is equal to

$$< n_{reset}^2 >= \sqrt{\frac{kT}{C_{pd}}} \tag{3.30}$$

where C_{pd} is the photodiode capacitance, *k* is the Boltzmann constant and *T* is the absolute temperature. This noise is also often refereed to as *kT/C* noise. The reset noise can be effectively canceled using correlated double sampling (CDS) as described in section 3.5.2.

Thermal Noise - The thermal noise originate from the thermal agitation or random motion of the electrons within a resistance. It is also refereed to as Johnson Noise. The power spectral density of the thermal noise in a voltage representation is given by

$$< n_{thermal}^2 >= 4kTRB \tag{3.31}$$

where *k* is Boltzmann constant, *T* is the absolute temperature, *R* is the resistance and *B* is the noise equivalent bandwidth. In many cases noise equivalent bandwidth is taken to be 1Hz and thus equation (3.31) can be simply written as *4kTR*. The thermal noise is directly proportional to the temperature.

1/f or Flicker Noise - This occurs mainly due to fluctuation in the conductivity. The carrier can fluctuate due to the presence of $Si - SiO_2$ interface trap, mobility fluctuation or even both. The defects at the interface of the $Si - SiO_2$ introduces random current fluctuations due to trapping and detrapping of signal electrons. In CMOS image sensors, the *1/f* noise mainly appears from the source follower transistor [11]. The *1/f* noise generated inside a MOS transistor can be described by a voltage source in series with the channel, and the power spectral density is given by

$$< n_{1/f}^2 >= \frac{K}{C_{ox}WL} \cdot \frac{1}{f} \tag{3.32}$$

where *K* is a process dependent parameter, C_{ox} is the oxide capacitance, *W* and *L* are the width and length of the transistor and *f* is the frequency of operation. *1/f* noise is dominant at low frequencies but for high frequencies it drops below the thermal noise.

3.4.5.2 Pattern Noise

The pattern noise is a spatial variation of the signal from pixel to pixel. For CMOS image sensors, the components that contribute to the pattern noise are the offset and

the gain variation of the devices used in the pixel. The pattern noise can be again divided into two categories, Fixed pattern noise (FPN) in the dark characterized by dark signal non-uniformity (DRNU) and fixed pattern noise in light characterized by photoresponse non-uniformity (PRNU).

In dark, the main source of FPN is the offset variations. There are two main sources of the offset-caused FPN, the mismatch of in-pixel or column level transistors and the dark current generation non-uniformity inside the pixel. The dark current generated inside the pixel is a random process, which is different for different pixels. Thus DRNU is an offset between pixels in dark. DRNU can be corrected by subtracting a dark frame from the signal frame. The DRNU is always measured in dark.

The photo response non-uniformity (PRNU) describes the variations in sensitivity to illumination between pixels of the CMOS image sensor. The PRNU is mainly due to the variations in the detector dimensions, doping concentration, thickness of the the overlayers above the pixels and the spectral sensitivity of the photodiode. The PRNU is measured only in illumination, in the absence of illumination or at low illumination levels, the fixed pattern noise (FPN) dominates over PRNU. The PRNU becomes constant at very high illuminations regardless of the signal level.

The fixed pattern noise and the PRNU are always expressed as a percentage of the maximum signal output or saturation output.

3.5 CMOS Image Sensor Peripherals

The architecture of an conventional CMOS image sensor is shown in figure 3.8. The main regions of the image sensor are the pixel array, row and column selector or address decorder, column processing circuit and an output amplifier or Analog-to-Digital converter *(ADC)*. In the following section, the addressing scheme for CMOS image sensors, column processing circuits and output amplifier/ADC for CMOS image sensors are described.

3.5.1 Addressing

In CMOS image sensors, to address each pixel, a scanner or a decoder is used. A scanner consists of a latch array or shift register array to shift the the data according to a clock signal. The use of shift registers for the vertical and horizontal access, allows the pixels to be addressed sequentially. A address decoder allows pixel to be randomly selected and read, which is really an advantage of the CMOS image sensors. The address decoder uses logic gates to generate the address of the pixel to be addressed.

The image collection mode of an area array sensor can be either interlaced or progressive scan (non-interlaced). In the progressive scan, each row is first selected and then each pixel in the given row is processed. This is more sequential in nature. In the interlace scan, the array of the sensor is divided into even and odd fields. The odd fields consists of all odd numbered rows, and the even field consists of all the

even numbered rows. Initially half of the frame (odd or even field) is recorded and then the remaining frame is read, the two half frames are then combined to form the final image. The interlaced technique is used by the major broadcast standards PAL and NTSC and is done to reduce the bandwidth of the image for transmission. A major drawback of the interlaced method is encountered when the subject is moving. Since the two fields are separated in time, the moving object results in blurring. The progressive scan reads the entire image in one go, thus reduces the blur effects when capturing images of moving objects.

3.5.2 Column Processing Circuits

The column processing circuit contains a correlated double sampling circuit. The correlated double sampling is often used to reduce the fixed pattern noise in the pixel. A sample CDS circuit is shown in figure 3.17. The CDS circuit includes two memories to store the reset and the sample signal, and two switches allowing access to the two memories. When the switch S_{SHR} is switched ON, the reset sample from the pixel is sampled onto the capacitor C_{reset}. Along with reset voltage, the reset noise is also sampled on the capacitor C_{reset}. The total voltage sampled on C_{reset} during sampling reset can be written as

$$V_{SHR} = V_{RST} + V_{offset} + V_{rn} \tag{3.33}$$

where V_{SHR} is the voltage sampled on the reset sampling capacitor, V_{RST} is the photodiode reset voltage and V_{offst} is the pixel off-set and V_{rn} is the reset noise sampled on the photodiode capacitance. After the end of the exposure period, the signal is then sampled onto the capacitor C_{signal} when the switch S_{SHS} is switched ON. Along with the signal, the resent noise is also sampled on to the capacitor C_{signal}. The sampled reset noise is the same as the reset noise sampled on C_{reset}, as there have been no reset in between sampling the reset voltage and the signal voltage. The total voltage sampled on C_{signal} during sampling signal can be written as

$$V_{SHS} = V_{SIG} + V_{offset} + V_{rn} \tag{3.34}$$

where V_{SHS} is the output voltage sampled on the signal sampling capacitor, V_{SIG} is the photodiode signal voltage and V_{offst} is the pixel off-set and V_{rn} is the reset noise sampled on the photodiode capacitance. A differential amplifier is used to subtract the sampled reset voltage from the sampled signal voltage, to obtain the pixel output proportional to the light intensity. Subtracting equations (3.33) and (3.34), we get the differential voltage as

$$V_{DIFF} = V_{SIG} - V_{RST} \tag{3.35}$$

From equation (3.35), it can be observed that the reset noise and the pixel off-sets completely gets canceled. The output voltage is free from reset noise and thus the

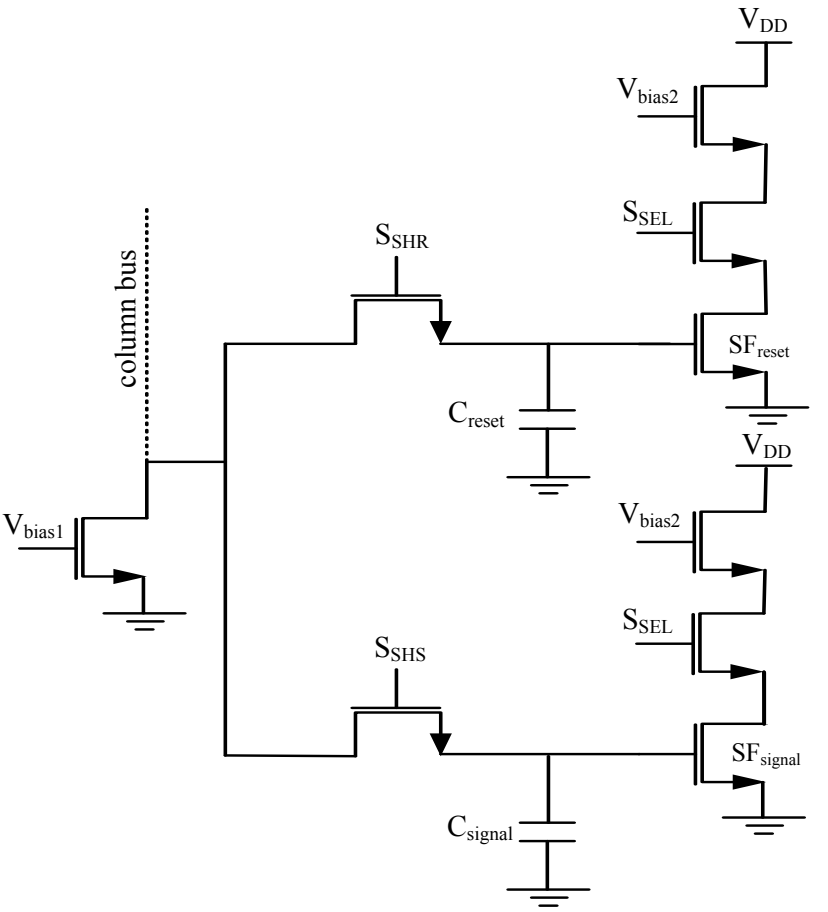


Fig. 3.17 Correlated double sampling

CDS operation helps in reducing the fixed pattern noise of the pixel. The sampled values are buffered through the source followers SF_{reset} and SF_{signal} to the output amplifiers.

A multiple reset, during the CDS operation makes the reset noise uncorrelated. If CDS is performed on the uncorrelated noise, the resulting noise power increases instead of decreasing. The reset process of a 3-Transistor pixel thus becomes very important. There are two ways to reset the photodiode, "soft reset" and "hard reset". In the soft reset, the gate and drain potential of the reset transistor are the same. When the reset transistor is turned ON, the photodiode node voltage increases to $V_{DD} - V_{TH}$, where V_{TH} is the threshold voltage of the reset transistor. When the node potential reaches $V_{DD} - V_{TH}$, the gate-to-source potential V_{GS} of the reset transistor is not sufficient to keep the reset transistor into saturation. The reset transistor thus starts to operate in weak inversion. The weak inversion operation

introduces non-linearity, however since the electron flow in the transistor is through diffusion and unilateral, the reset noise power in voltage square is actually $kT/2C$ which is less than kT/C [12].

In hard reset, the gate potential is held higher than the drain potential of the reset transistor. This makes sure that the gate-to-source potential V_{GS} of the reset transistor is always sufficient to hold the reset transistor in saturation during reset. The photodiode potential node is reset to a fixed voltage and thus the non-linearity in the photoresponse is removed. However as the reset transistor is operating in saturation, there exists a bilateral way for the electrons to flow (from photodiode node to the power supply and the power supply to the photodiode node), the reset noise power in voltage square is kT/C. Thus the hard reset introduces higher reset noise compared to soft reset.

3.5.3 Analog-to-Digital Conversion

The analog signal from the column can be converted into a digital signal using an analog-to-digital converter. The ADC can be chip-level, which is common for all the pixels as shown in the figure 3.8. Most of the early camera-on-a-chip products made in CMOS were equipped with chip level ADC. The chip-level ADC assured the uniformity of processing of the signal from the pixels. However the chip-level ADCs has two potential disadvantages. Firstly, the ADC have to operate at a high sped to be able to readout the complete imaging array. Second the analog signal chain with the chip-level ADC is longer compared to the column-level (figure 3.18) or pixel-level A/D conversion (figure 3.19). The longer signal chain path can have more noise sources and thus affect the overall performance.

The ADC can also be placed in the column processing circuit with each column having its own ADC or can be incorporated inside the pixel. The column level ADCs are really helpful when the number of pixels in the pixel array is very large. The column ADCs can operate in parallel and thus high speed can be obtained compared to the single chip level ADC. In the column-level ADCs, the correlated double sampling and the A/D conversion are now performed at the column, and thus the pixels output can be digitized concurrently. The analog signal processing chain is also shorter than that of an imager with a chip-level ADC. There are two main drawbacks of column-level ADCs. Firstly the column level non-uniformities needs to be corrected using additional circuity. Secondly, because of the parallelization, imagers with column-level ADC usually requires large area compared with the chip-level ADC.

The pixel-level ADC has the shortest analog signal chain path. The pixels with ADC are also called as digital pixel sensor (DPS) described in section 3.3.3. The pixel with ADC can achieve very high readout speeds. However the pixel level ADCs will introduce non-uniformities which needs to be compensated. The main drawback of the pixel-level ADC is that it requires lots of in-pixel circuitry, which reduces the fill factor of the sensor.

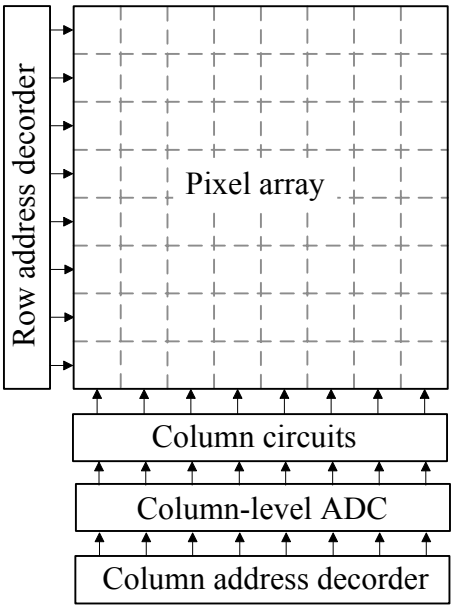


Fig. 3.18 Column-level ADC implementation

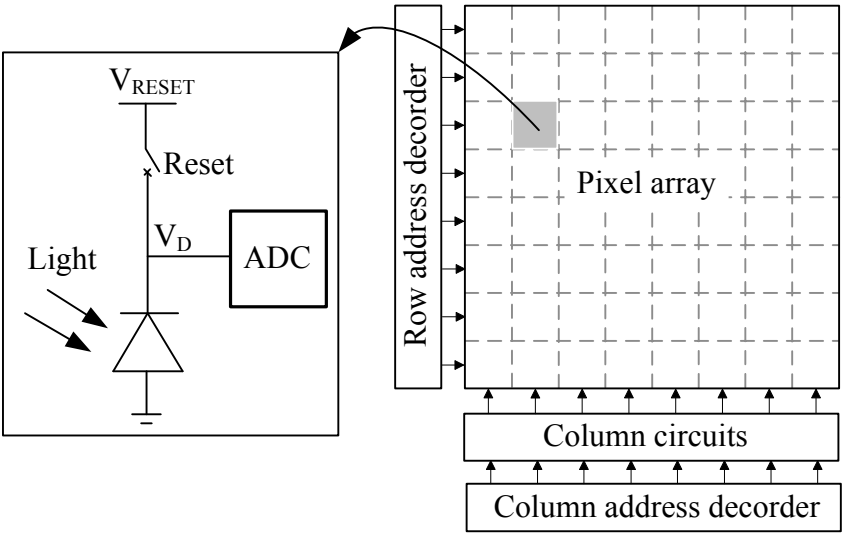


Fig. 3.19 Pixel-level ADC implementation

3.6 Designed Sensor Overview

The pinned photodiode described in section 2.4 is used to design a prototype image sensor using 0.18μm 3-metal-1-poly CIS process from UMC. The sensor architecture of the designed image sensor is shown in figure 3.20.

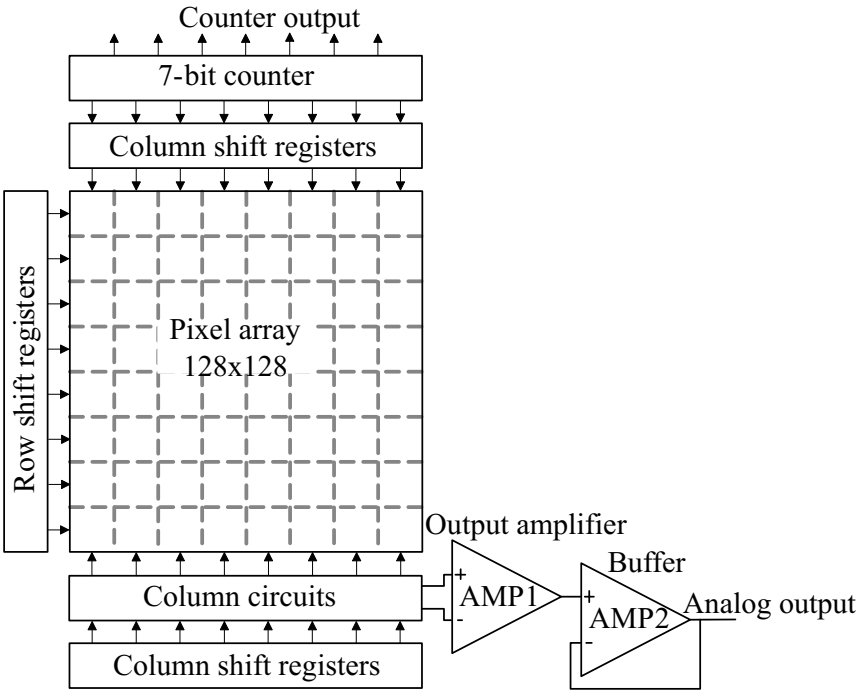


Fig. 3.20 Sensor architecture

The chip is divided into four main blocks:

1. the pixel array,
2. the analog readout circuit,
3. the digital readout circuit and
4. the row select logic plus timing control block.

The pixel array is located at the center, with the photodiodes and their associated circuitry for analog and digital computations. It contains an array of 128 by 128 pixels. Each pixel contains a pinned photodiode and 33 transistors including *TG* to perform low level image processing. The size of the pixel is 25μm x 25μm and the size of the photodiode is 10μm x 10μm which corresponds to a 16% pixel fill factor. One half of the pixel array of the designed image sensor is used to analyze polarization information using metallic wire grid structures explained in chapter 4. The other half is available for normal imaging applications.

The analog readout circuit is placed below the pixel array. It consists of column-wise double differential sampling circuits, an output amplifier, a buffer and a column shift register. The analog output provides the analog voltage at each pixel.

The digital readout circuit is placed at the top and consists of a 7-bit counter and a column shift register. The 7-bit counter is used to count the number of active high pixels in each row.

Finally, the left side is dedicated to the row select logic and timing control blocks to address each row of pixels sequentially.

Table 3.1 summarizes the specifications of the sensor.

Table 3.1 Sensor specifications

Process	0.18μm 1P3M UMC CIS process
Active imager size	3.2 mm(H) x 3.2 mm(V)
Chip Size	4 mm(H) x 5 mm(V)
Active pixels	128 x 128
Pixel size	25μm x 25μm
Shutter type	Global shutter
Modes of operation	Double differential sampling (DDS) Differential imaging (DI)
Maximum data rate/master clock	64 MPS / 32 MHz
Supply voltage	1.8V
Fill Factor	16 %
Pixel type	33T (Custom designed)
Target frame rate	30fps

The image sensor is operated in global shutter mode. In the global shutter operation, the entire imaging array is globally reset before integration and all rows start and stop integrating at the same time. A rolling shutter on the other hand has a start and stop time shifted for every row by one time unit with respect to the previous row. Global shuttering avoids image smear or distortion of fast-moving objects during readout as all the pixels are reset and integrated over the same interval. While rolling shutter is good for static or slow moving objects however they produces image smear and are incompatible with *CDS* (Correlated double sampling) processing.

3.6.1 Pixel Architecture

A pixel can either be a passive pixel *(PPS)* or an active pixel. A passive pixel is composed of a photodiode and an addressing transistor. The *PPS* is very simple and compact however it suffers from noise problems due to the mismatch of the small photodiode and the large column capacitance. In contrast to a passive pixel, an active pixel *(APS)* incorporates a buffer (or amplifier) in the pixel which solves the noise problems due to the large column capacitances. The pixel used in the design is a smart pixel which is a form of an active pixel sensor. Smart pixels are able to perform

parallel processing of large pixilated images, thus reducing the amount of signal sent off-chip. Although this type of pixel significantly improves the performance of the pixel, it severely limits the fill factor [13].

The desired functionalities of the compound eye of the insects namely, motion detection, high sensitivity to light intensity and polarization as described in section 1.2 are to be replicated in the pixel. The polarization is detected using the metals layers of standard CMOS and is described in detail in chapter 4. As discussed in the section 1.1, insects detect motion based on the change in state of its individual receptors creating a *"flickering effect"*. To reproduce this effect, in-pixel memories are needed to detect change in the state of the pixel by comparing current image with previous image. Further as the insects sees things as only "on" and "off", which digitally means '1' or '0', the stored images of the current and the previous image need to be binarized. Further to enhance the sensitivity to low light intensities, the dynamic range of the image sensor is extended using partial charge transfer mechanism. In partial charge transfer, the accumulated photoelectrons resulting from the incident illumination is sampled multiple times. The design choice to have in-pixel memories for detecting motion provides for storing the multiple samples obtained using partial charge transfer mechanism to enhance the sensitivity. In the following sub-sections the design of the pixel and the pixel operation is explained in detail.

Figure 3.21 shows the transistor level circuit diagram of the pixel. In the subsequent sections, a simplified version of the pixel architecture will be used to explain the pixel operation.

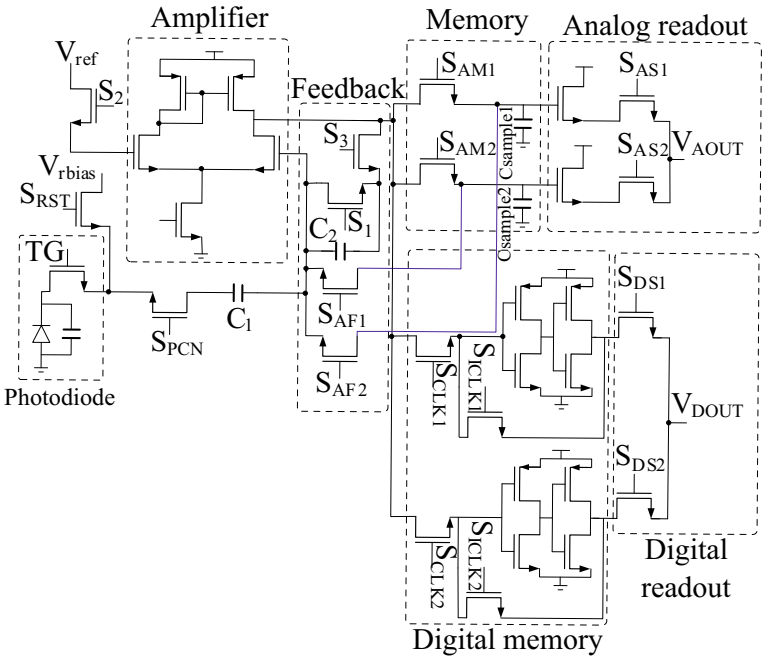


Fig. 3.21 Pixel schematic

3.6.2 *Pixel Operation*

The operation of the pixel can be divided into four phases: image capture, image storage, digitization and readout. The image capture phase includes the reset of the photodiode and integration of the available photons in the photodiode. The image storage phase transfers the signal integrated at the photodiode from the sense node to the available analog memory. The digitization phase digitizes the stored analog signal. During the readout phase the data is transferred to the analog and digital signal processors via the respective column buses.

3.6.2.1 Image Capture Phase

The light sensing part used for the pixel is a $p^+/n/p$ pinned photodiode. The cross section of the pinned photodiode is shown in the figure 3.22, together with the transfer gate *(TX)* and the reset switch (S_{RST}).

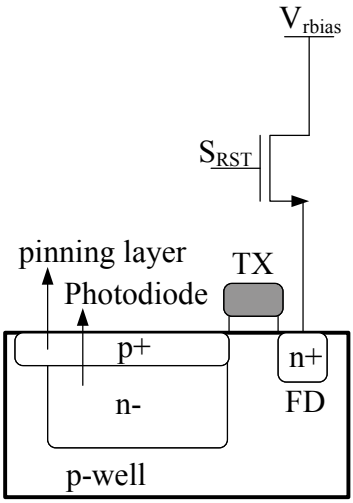


Fig. 3.22 Pinned photodiode with reset

The image capture phase begins with a reset of the pixel by closing the (S_{RST}) switch. The voltage at the floating diffusion node *(FD)* is set to the reset voltage V_{rbias}. To ensure that the reset transistor switches properly, the voltage V_{RST} at the gate of the reset transistor needs to be a threshold voltage (V_{TH}) above V_{rbias}. This is also known as "hard reset". Another form of reset is the "soft reset" where the gate and the drain potential of the reset transistor are at the same level. Soft reset has a noise advantage over hard reset because the electrons in soft reset can only move uni-directionally from the photodiode to the reset transistor compared to the bi-directional motion of the electrons in hard reset. However a soft reset produces a higher image lag and also decreases the output voltage swing possible at the FD node thus hard reset is used in this design.

For the hard reset, V_{rbias} can be expressed as

$$V_{rbias} = -V_{RST} - (V_{TH} + V_{ovd}) \tag{3.36}$$

where V_{TH} is the threshold voltage of the reset transistor (including body effect) and V_{ovd} is its overdrive voltage. A *NMOS* transistor is used as the reset switch instead of a *PMOS* transistor in this design as a *PMOS* transistor needs an additional *n*-well, which would not only increase the pixel area but would also contribute to the noise by the additional electrons generated by the captured photons in the *n*-well. The reset transistor is a thick oxide transistor capable of handling 3.3V provided by the 0.18μm CIS technology. The voltage V_{RST} is set to 3.3V and V_{rbias} at 1.8V as $V_{TH} + V_{ovd}$ would not exceed 1.5V for the chosen technology. The reset transistor is designed with minimum dimensions ensuring that there are few charges in the channel when it is conducting, and reducing the charge injection when it is turned off. When S_{RST} is closed (connected) the transfer gate *TX* is also closed, to effectively reset the photodiode to its pinning voltage. For the designed photodiode the pinning voltage is at 1V. The transfer gate has a channel width of 10μm and a channel length of 1μm. The wider transfer gate helps the charge transfer from the photodiode to the floating diffusion node. The floating diffusion capacitance must be large enough to hold all the charges transferred from the photodiode. However, it cannot be made too large as it decreases the conversion gain.

When the transfer gate opens, the photodiode starts accumulating the photo-generated charge. The time spent accumulating charge is referred to as integration time or exposure period. At the end of the integration time, the accumulated charge is transferred to the *FD* by closing the transfer gate *TX*. Figure 3.23 shows the potential diagram of the pinned photodiode for complete charge transfer. When the photodiode is fully depleted, after the charge transfer it returns to its pinning voltage as shown in figure 3.23(a). After the photodiode is reset to its pinning voltage, charges start to accumulate on the photodiode as shown in figure 3.23(b). The amount of charge collected is directly proportional to the illumination level till the photodiode reaches saturation when the full well is reached.

At the end of the integration period, the transfer gate *TX* is closed, which allows the transfer of accumulated charges from the photodiode into the floating diffusion node as shown in figure 3.23(c) and (d).

Besides a complete charge transfer, the photodiode can also transfer accumulated charges partially after each exposure period. This is known as partial charge transfer mode of operation, where the total exposure period is composed of multiple shorter exposure periods. After each small exposure the accumulated charges in the photodiode are transferred to the *FD* node. The operation of the pixel in the partial charge transfer mode when sampled twice is shown in figure 3.24.

Figure 3.24(a) shows the depleted state of the photodiode. The charge is accumulated (figure 3.24(b)), and after a certain accumulation time (T_1) the transfer gate is closed (figure 3.24(c)) transferring the accumulated charge into the floating diffusion (figure 3.24(d)). After the first transfer, the photodiode continues to

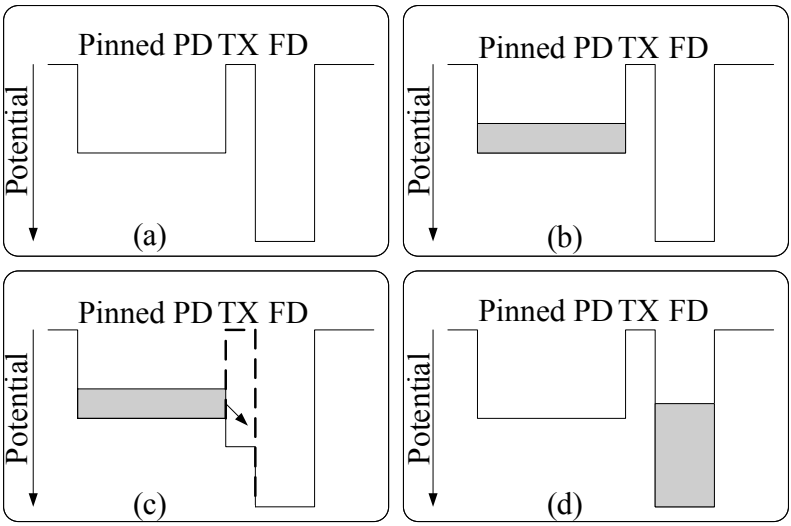


Fig. 3.23 Pinned photodiode operation a) before charge integration, b) during and after charge integration, c) during charge transfer, d) after charge transfer

accumulate charge until the end of the exposure time (T_{FT}). The accumulated charge is transferred to the floating diffusion by closing the transfer gate as shown in the figure 3.24(e)-(h). Thus the charges are transferred to the *FD* at two time instances: (T_1) and (T_{FT}).

3.6.2.2 Image Storage Phase

The simplified circuit for the image storage phase along with the corresponding timing diagram is shown in the figure 3.25. In the image storage phase, the voltage change from the reset level at the *FD* node due to the transferred photo-charge is sampled onto one of the two available sampling capacitors C_{sample} when the switch S_{AM} is closed in figure 3.25.

The amplifier *AMP* used is a simple differential amplifier. It serves two functions: amplification and comparison. When S_3 is closed, the amplifier works in closed loop configuration. When the analog signal voltages at the sense node are sampled onto the analog memories, the amplifier amplifies the signal at the sense node by the closed loop gain provided by the capacitors C_1 and C_2. The pixel is designed with a closed loop gain (A_{PIXEL}) of 7. The switch S_2 connects the globally distributed reference voltage V_{ref} to the non-inverting terminal of *AMP* and the switch S_{PCN} is used to connect or disconnect the pinned photodiode to the the *AMP*.

Autozeroing is used to reduce the effects of the input-referred dc offset voltage (V_{off}). When S_1 and S_3 are closed, the amplifier is connected as a voltage follower and the offset voltage V_{off} is stored on the capacitor C_1. The charge equations when S_1 and S_3 are closed can be written as

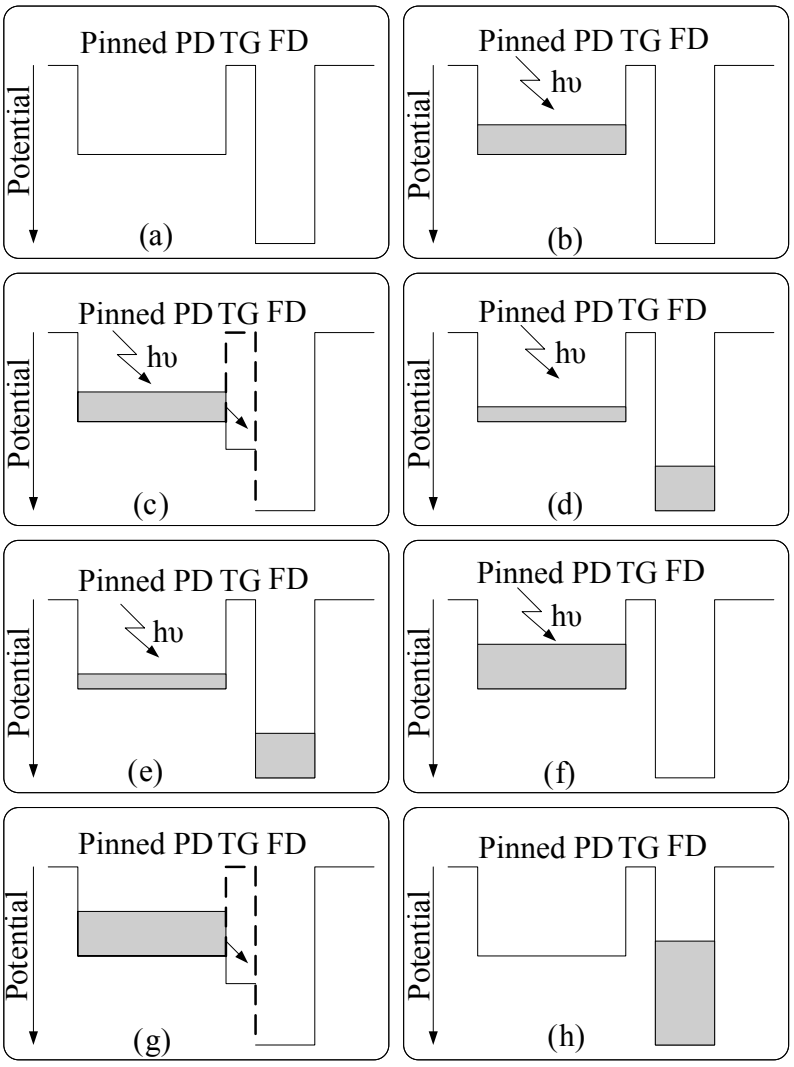


Fig. 3.24 Pinned photodiode operation in partial charge transfer mode

$$Q_1 = C_1(V_{ref} + V_{off} - V_{rbias}) + C_2 V_{off} \tag{3.37}$$

where Q_1 is the offset charge, V_{ref} is the reference voltage applied to the non-inverting terminal of *AMP*, V_{off} is the offset voltage, V_{rbias} is the reset voltage and C_1 and C_2 are the capacitors. When the signal is sampled, switch S_1 is opened and the charge equation can be written as

$$Q_2 = C_1(V_{ref} + V_{off} - V_{sig}) + C_2(V_{ref} + V_{off} - V_{out}) \tag{3.38}$$

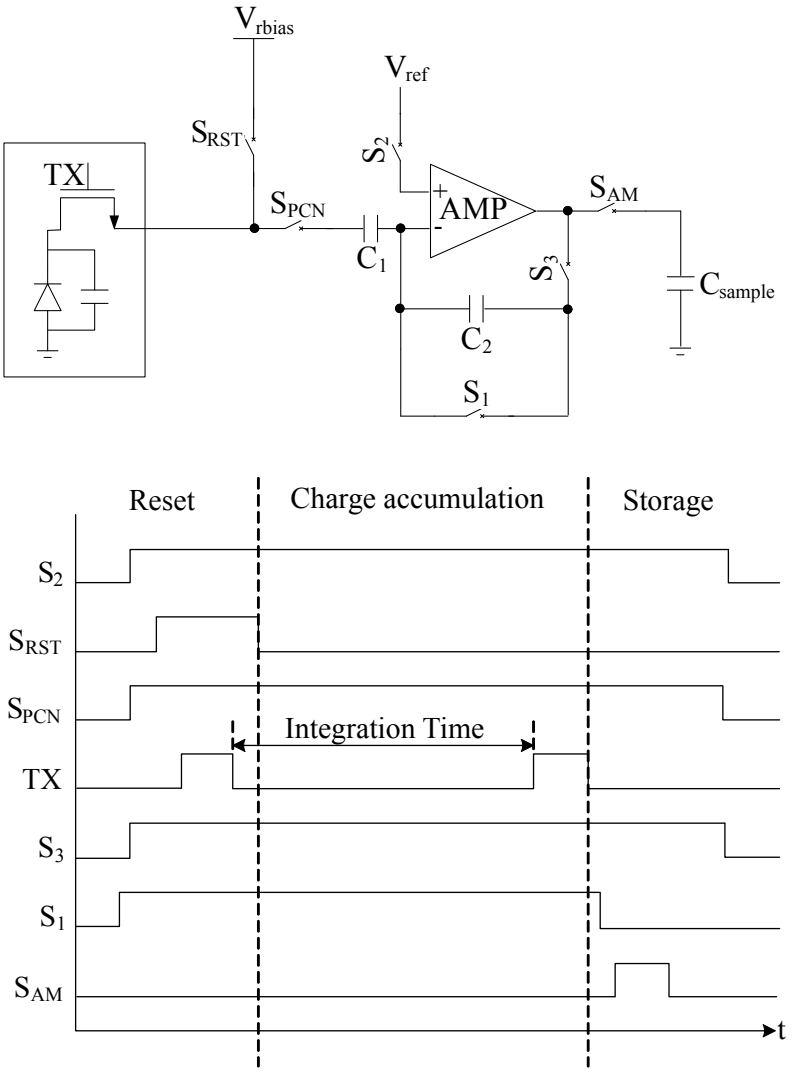


Fig. 3.25 Pixel in image storage phase and associated timing diagram

Because the change in charge in the capacitors is reflected in the output voltage V_{out}, the output voltage can be derived from equation (3.37) and (3.38) and is free of offset voltage of the amplifier.

$$V_{out} = \frac{C_1}{C_2}(V_{rbias} V_{sig}) + V_{ref} \tag{3.39}$$

The designed image sensor can be operated in two modes: double differential mode *(DDS)* and differential imaging mode *(DI)*. In case of *DDS* each pixel samples the voltage at the *FD* node twice: once after reset and again after the exposure period and it stores the sampled values in the two available analog memories. In the differential mode two samples are stored in the analog memories after different exposure periods. The *DDS* mode of operation is further elaborated in section 2.5.5.1 while the *DI* mode of operation is explained in chapter 7.

3.6.2.3 Image Binarization Phase

The simplified schematic of the pixel in the image binarization phase and its associated timing diagram is shown in the figure 3.26.

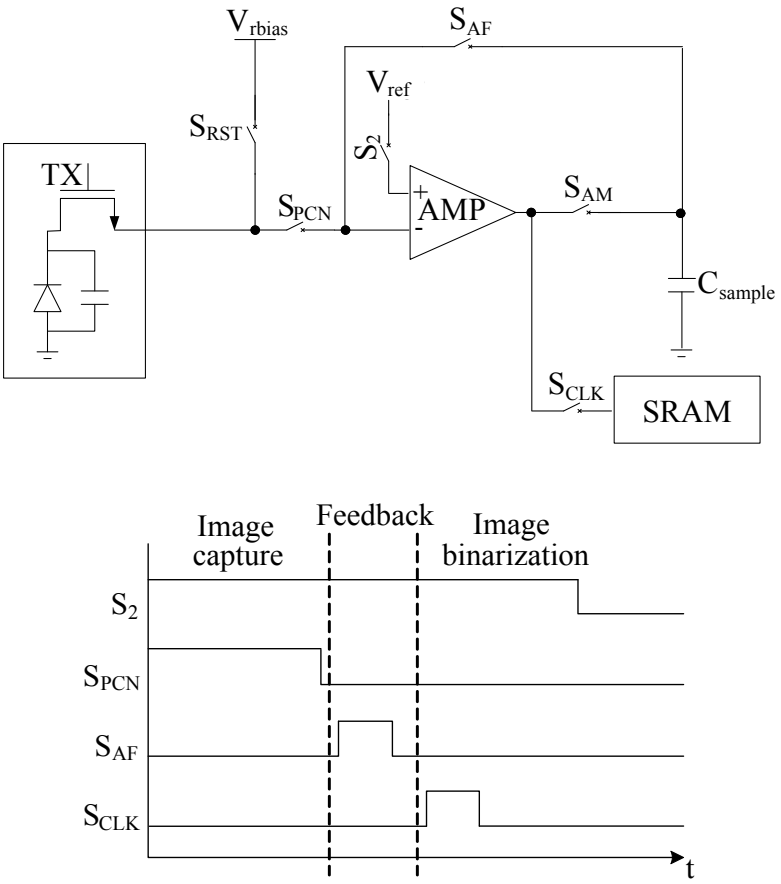


Fig. 3.26 Pixel in image storage phase and associated timing diagram

After an image is captured, the photodiode is disconnected from the processing elements by opening the switch S_{PCN}. The analog signal sampled on C_{sample} is compared with a fixed reference voltage V_{ref} when the switch S_{AF} is closed, by using the amplifier as a comparator. The globally distributed reference voltage V_{ref}, is connected to each pixel's comparator non-inverting input while the feedback signal stored in the analog memory is connected to the inverting input of the comparator. The resulting binary output voltage is stored in the *6T-SRAM* cell when the switch S_{CLK} is closed. Two such *SRAM* cells are available in each pixel to store the binary data of the two sampling capacitors. Writing into the memory is locally controlled by the comparator. The switch S_{AM} is always open in the binarization phase.

3.6.2.4 Image Readout Phase

The final phase of the pixel operation is the readout phase when each pixel is read out sequentially. The source follower *SF* loads the column bus V_{AOUT} via the analog row selection switch S_{AS} with the voltage sampled on the sampling capacitor during readout of the pixel as shown in figure 3.27. The switch S_{DS} loads the column bus V_{DOUT} with the binary value stored in the *SRAM* cell. The control signals for switches S_{AS} and S_{DS} are generated by the row addressing logic explained in section 2.5.4.

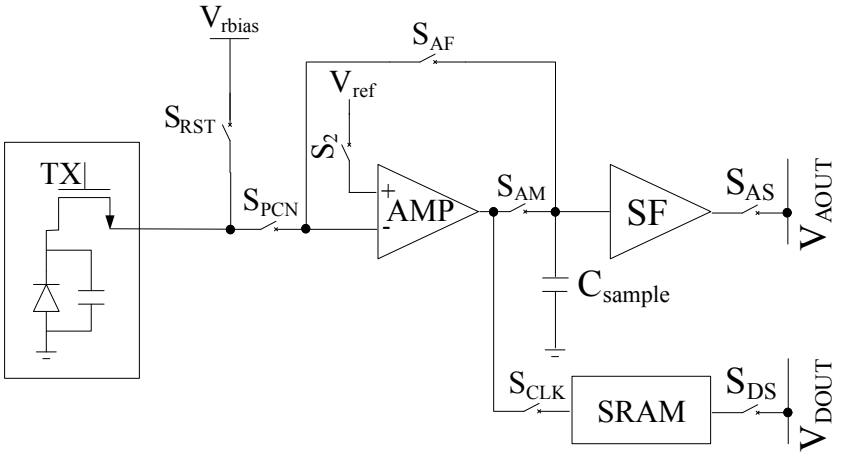


Fig. 3.27 Pixel (Image) readout phase

3.6.3 Pixel Layout

Figure 3.28 shows the layout of a single pixel block. The large square at the center is the pinned photodiode with a dimension of 10μm by 10μm. The analog memories C_{sample} are implemented with Metal-Metal (MiM) capacitors and are placed below the photodiode in figure 3.28. The amplifier/comparator is placed on the left hand side while the digital memory is placed on the right hand side.

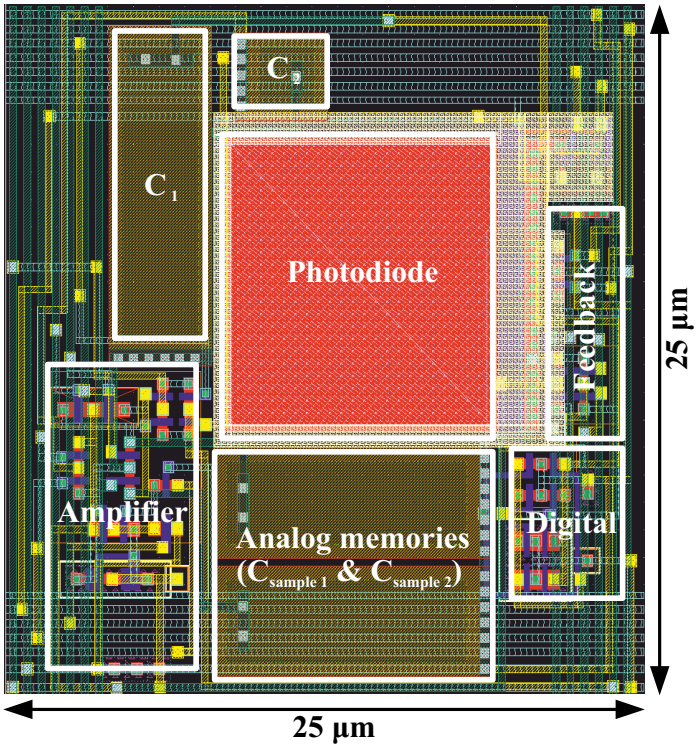


Fig. 3.28 Pixel schematic and layout

3.6.4 Row/Column Addressing Logic

Several architectures for pixel addressing are available in CMOS image sensors. The row select logic used in the designed image sensor is column parallel readout architecture. Column parallel readout architecture for a 128 x 128 pixel array is shown in figure 3.29. The row address block generates the row select pulse to select a row and the column address pulse generates the column select pulse to select the column in the selected row to be read.

The column parallel readout is progressive in nature. Progressive scan readout of an area format CMOS image sensor involves selecting a row and simultaneously reading and processing all pixels in the selected row. Column readout is sequential in nature for the designed imager, where each pixel is readout serially. The progressive scan readout sequence is shown in the figure 3.30.

In image sensors, two readout control structures are commonly used to generate the control signals for selecting the columns and the rows: shift registers and decoders. The advantage of using shift registers over decoders is that shift registers are easy to implement and use fewer transistors. However random access to pixels is not possible with shift registers as they produce only sequential output from the first element to the last one. Decoders are preferred for random access readout.

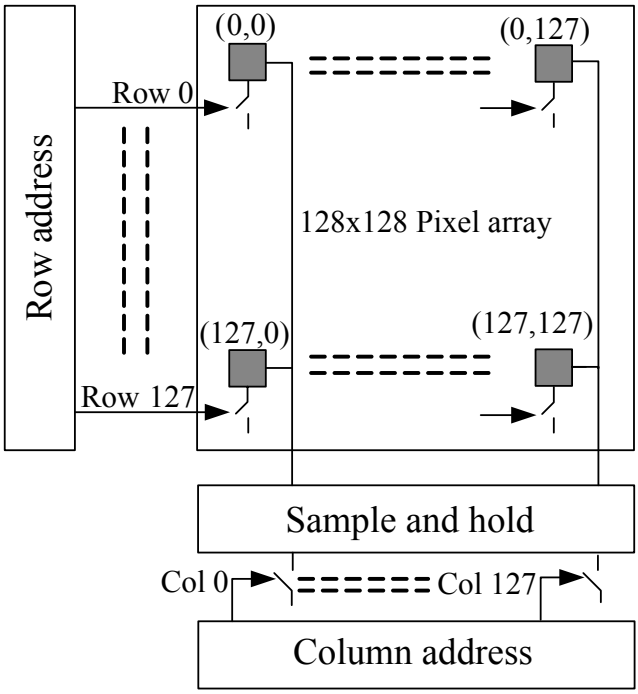


Fig. 3.29 Column parallel readout architecture

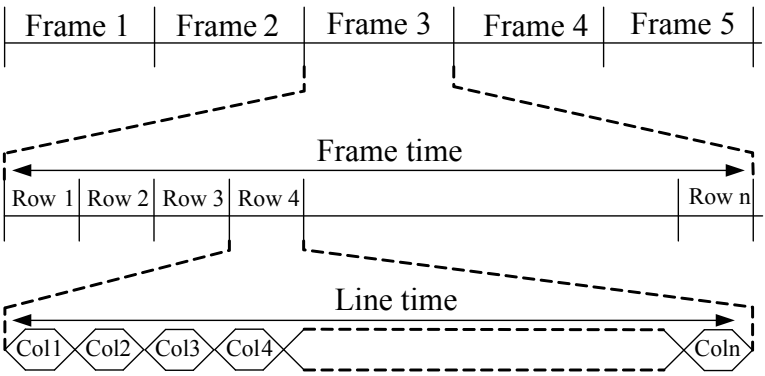


Fig. 3.30 Progressive scan readout

However in this design shift registers are preferred for its simplicity in implementation. In this design, the row selection logic circuit is composed of a non-overlapping clock generator block and a shift register block, which is a cascade of shift registers. Figure 3.31 shows the implemented shift register.

The shift register is a serial-in serial-out type. The data is delayed by one clock cycle at each stage. *PHI1* and *PHI2* are non-overlapping clocks. When the clock

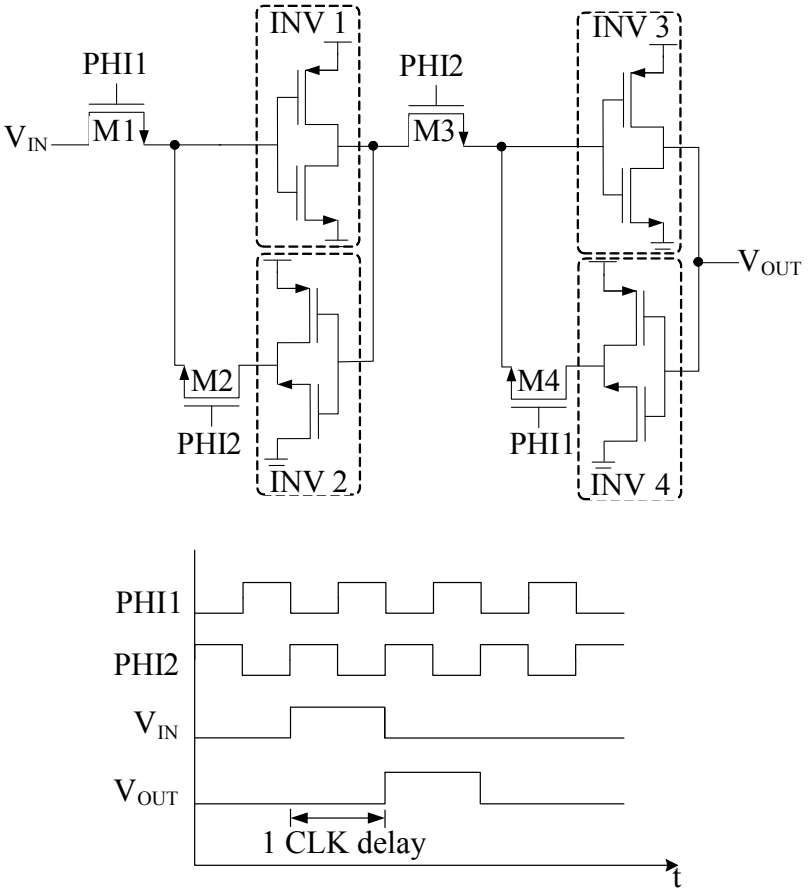


Fig. 3.31 Shift register and associated clock timing

PHI1 is high the input signal to the shift register V_{IN} is saved in the *INV1* and *INV2* loop. This saved signal is shifted out as an output V_{OUT} when *PHI2* is high via the *INV3* and *INV4* loop.

The non-overlapping clock signals *PHI1* and *PHI2* are generated on-chip using the non-overlapping clock generation circuit shown in the figure 3.32. The non-overlapping clock generation circuit is implemented with *NAND* gates and inverters. The non overlapping clocks *PHI1* and *PHI2* are fed to the shift register block to shift the row read control signal ROW_RD_SIGNAL to generate sequential outputs to read one row at a time. Both the analog signal processor and the digital signal processor in figure 3.20 have individual column select pulses generation circuit to select one column at a time. These circuits are similar to those used in the row select logic.

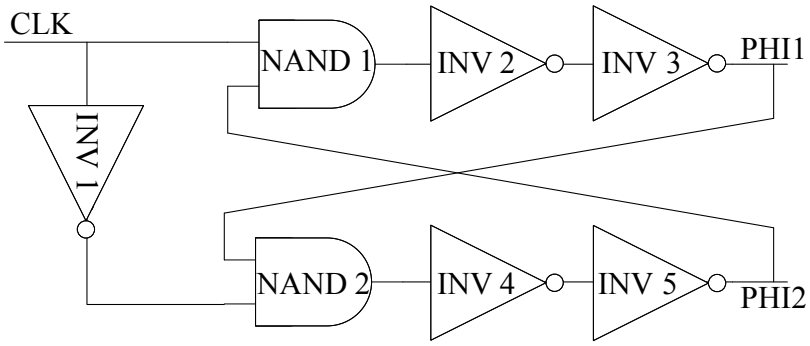


Fig. 3.32 Non-overlapping clock generation circuit

3.6.5 Analog Signal Chain

The simplified analog signal chain is shown in the figure 3.33. In the analog signal chain the analog signals stored in the analog memories $C_{sample1}$ and $C_{sample2}$ in the pixel are transferred to the sample and hold capacitors C_1 and C_2 in the column differential sampling circuit when the pixel is read. The stored signals in the sample and hold capacitors are buffered to the output amplifier *A3*. The differential output generated by *A3* is then buffered to the chip output pad using the voltage follower *A4*.

3.6.5.1 Delta Differential Sampling *(DDS)*

Delta differential sampling is based on using two successive track-and-hold *(T/H)* operations to reduce the low frequency noise and fixed pattern noise inherent in standard CMOS image sensor outputs. *DDS* reduces noise by sampling each pixel twice, once immediately after reset and again after exposure. The array is readout row wise, where the reset and the signal levels in the in-pixel memories are transferred to the column circuits one row at a time. However since all the pixels are reset at the same time the imaging array as a whole has a global shutter. By storing the pixel value after reset in a track-and-hold circuit, the pixel level fixed pattern noise can be subtracted from the actual pixel signal at exposure for a very accurate measurement.

The schematic of the sample and hold circuitry used in the designed image sensor is shown in figure 3.34. The signal on the pixel analog column bus V_{AOUT} is sampled to C_1 when both the S_{SHR} switch and the row select switch S_{AS1} are closed and to C_2 when the S_{SHS} switch and the row select switch S_{AS2} are closed. The crowbar switch S_{CAL} is opened when the pixel output line is sampled on the sample and hold capacitors.

The bottom plates of the sample and hold capacitors are connected to the externally generated reference voltages V_{REF_R} and V_{REF_S} to reduce the substrate noise [14]. As the bottom plates of the sample and hold capacitor sees a fixed reference voltage, the substrate noise through the the parasitic capacitor between

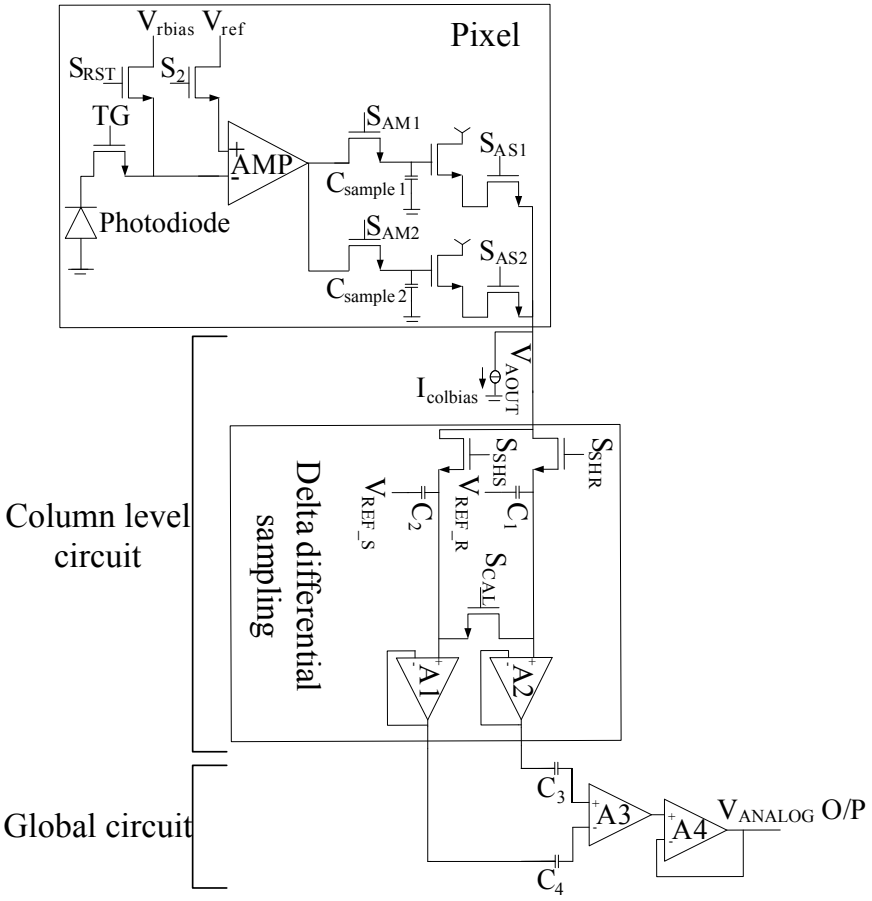


Fig. 3.33 Analog signal chain

the bottom plate and the substrate is of little importance. The voltage values used for V_{REF_R} and V_{REF_S} is 500mV, to provide enough bias voltage to the voltage followers *A1* and *A2*. The crowbar switch S_{CAL} shorts the two storage capacitors to compensate for the offsets of the voltage followers *A1* and *A2*.The timing diagram for the differential sampling operation and the S_{CAL} switch clock timing to read one row is shown in the figure 3.35.

The minimum value of the sample and hold capacitors are determined by *kTC* noise limitations. A 500fF hold and sample Metal-Metal capacitor is chosen, which gives a noise of 90μV. The output of the column sample and hold capacitors are fed into the output amplifier through the voltage followers *A1* and *A2*. The voltage followers *A1* and *A2* are similar to the *OTA* used in the pixel to shorten the design time.

The crowbar operation is initiated when S_{CAL} is closed. After the sampling of the column output on the sample and hold capacitor, the S_{CAL} switch is closed

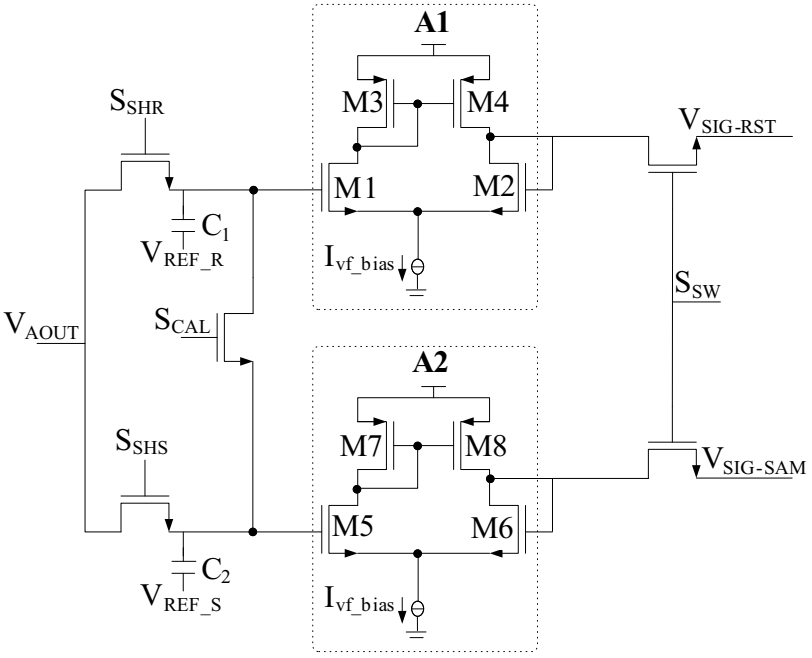


Fig. 3.34 Analog signal chain

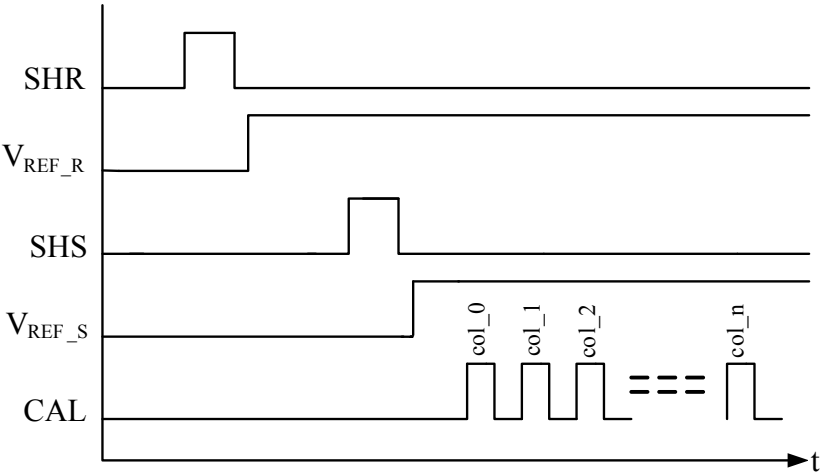


Fig. 3.35 DDS timing diagram

shorting the two capacitors. The output of the differential output amplifier changes depending on the voltage difference sampled on the sample and hold capacitors. The switch S_{SW} is used to transfer the voltages from the delta differential circuit to the output amplifier circuit.

3.6.5.2 Output Amplifier

The overall block diagram of the output differential amplifier *A3* is shown in figure 3.36.The signals V_{SIG_SAM} and V_{SIG_RST} are the signals in the sample and hold capacitors of the delta differential sampling circuit. The signals are stored on the left plates of the capacitors C_3 and C_4 respectively as shown in figure 3.36. Before the differential signals V_{SIG_SAM} and V_{SIG_RST} are amplified,

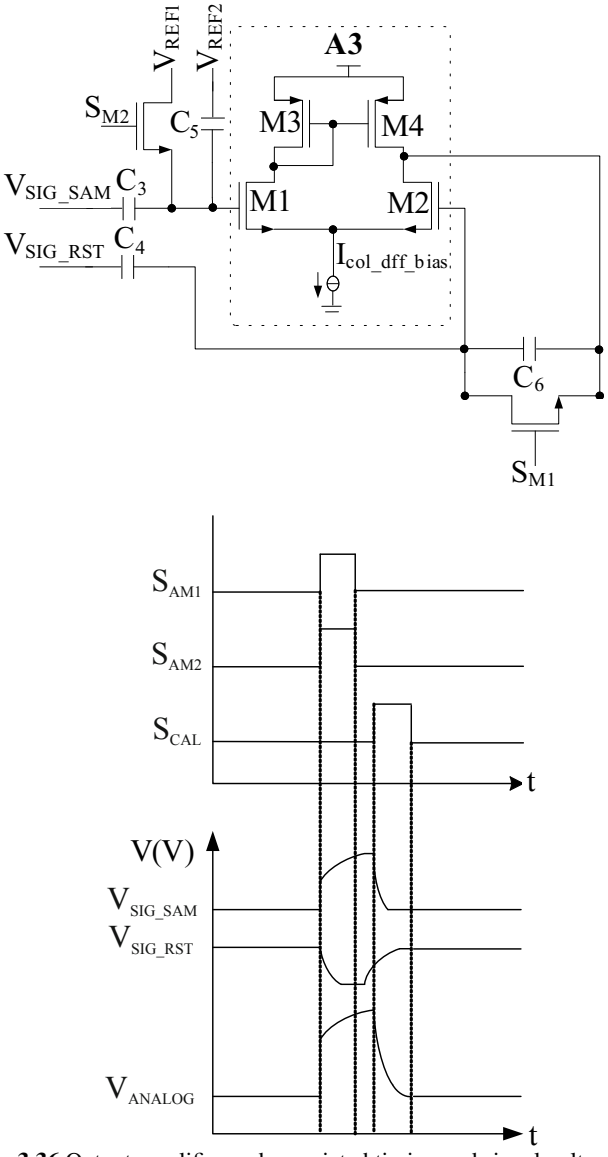


Fig. 3.36 Output amplifier and associated timing and signal voltages

the amplifier *A3* is connected in autozero mode to store the offset on the top plate of the capacitor C_6. Figure 3.36 also shows the clocking diagram of the switches S_{M1}, S_{M2} and S_{CAL} (from delta differential circuit block) and the corresponding variations in the signals in the sample and hold channels along with the obtained V_{ANALOG} output. The S_{M1} and S_{M2} transistors have minimum channel dimensions to reduce the charge injection when the switches are closed.

The amplifier A3 is a simple differential *OTA* with an *N*-type input stage. The values of C_3 and C_4 are 250fF. The feedback capacitor C_6 sets the voltage gain. C_6 is chosen to be 125fF, to provide for a closed loop gain (A_{OUTPUT}) of 2. A 125fF capacitor C_5 couples the DC reference voltage V_{REF2}. The reference voltages V_{REF1} and V_{REF2} are generated externally.

Table 3.2 Design values for figure 3.36

C_3, C_4	250fF
C_5, C_6	125fF
Gain (A_{OUTPUT})	2

3.6.6 Digital Signal Chain

The simplified digital signal chain is shown in figure 3.37. The binary data from the two *SRAM* cells is loaded onto the digital column bus *DOUT*. The tri-state buffer is used to invert the incoming digital signal and the *RS* flip flop synchronizes the tri-state buffer output for the counter. A 7-bit counter counts the number of active high pixels in each row. The outputs of the counter and the digital output are buffered out to the chip output pins.

3.6.6.1 Digital Counter

The designed counter is a synchronous up counter with an reset input. The reset option allows resetting the counter to an all-zero state when desired. The 7-bit digital counter is implemented using 7 modified *J-K* flip flops with their input shorted for toggle operation and additional *AND* gates as shown in the figure 3.38. The outputs of the counter Q_0, Q_1, Q_2, Q_3, Q_4, Q_5 and Q_6 are available for external interface to the FPGA for digital processing.

The *J-K* flip flop has been designed following a master/slave approach. Another alternative could be edge sensitive flip flops. However for the edge sensitive flip flops to work properly sharp clock edges are required and the threshold voltage of inverters and other gates be adjusted properly. Also the delay time through the two inverters at the input is critical. For these problems with clock and timing requirements, edge-sensitive flip-flops are not used in the design. The *J-K* flip flop can easily be toggled and thus is used in the counter in our design.

The output of the counter is all zero when the RST_CN (*RST* to the *J-K* flip flop) is held at logic high.

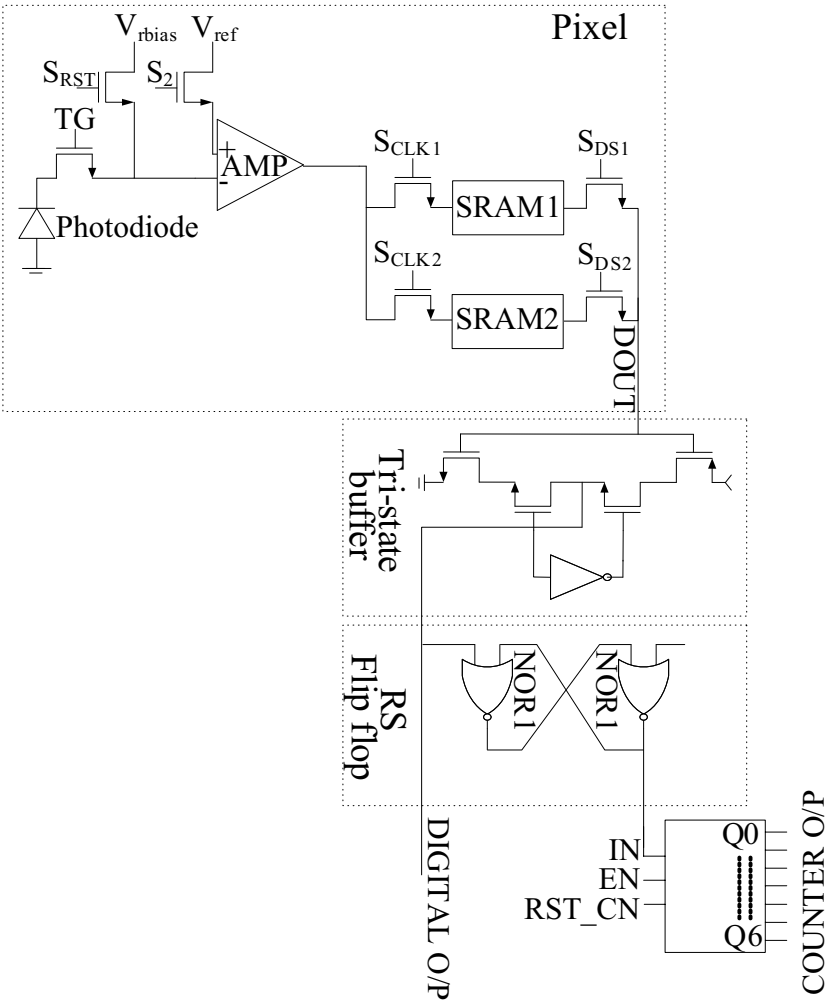


Fig. 3.37 Digital signal chain

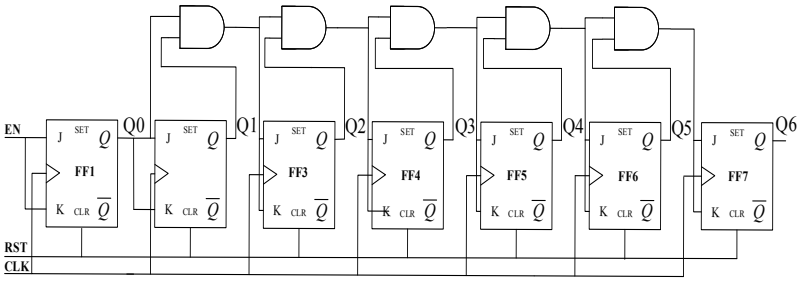


Fig. 3.38 Digital counter

The pixels in the column are read out sequentially. When the digital output from each pixel in the column selected by the column select signal is high and the RS_CLK of the *RS* flip flop is low then the output of the *RS* flip flop is set high which increments the counter.

3.7 Image Sensor Characterization

A chip micro-photograph of the fabricated image sensor is shown in figure 3.39. The chip was manufactured in a 0.18μm 1-poly 3-metals CIS CMOS image sensor technology. The imaging array has 128x128 pixels with a pixel pitch of 25μm. The chip size is 5mm x 4mm. The device is enclosed in a 68-pin JLCC (J Leaded Chip Carrier) package with a size of 25mm x 25mm. The chip also contains test structures which are used for testing pixels with different shapes of pinned photodiodes.

The characterization parameters for an image sensor are related to the parameters of the image quality and also the intended application. The parameters of interest for our application are the photoconversion characteristics, the sensitivity, the photon transfer curve to determine the conversion gain, and the dark current.

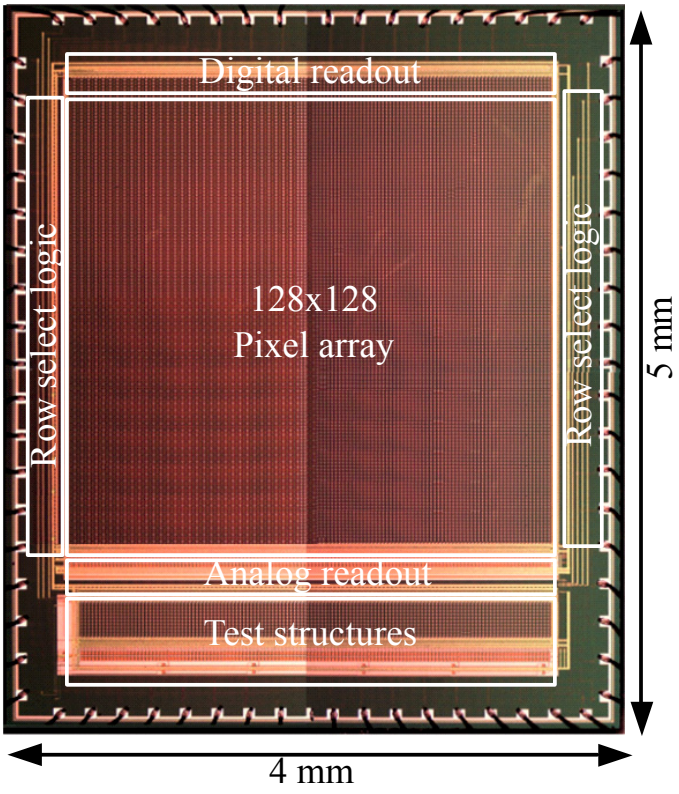


Fig. 3.39 Die photograph of the sensor

3.7.1 Test Setup

Figure 3.40 shows the experimental platform used to characterize the image sensor. Two *Xilinx Virtex-II* pro development boards [15] are used to generate clocks and control signals for the image sensor test board, *ADC* and frame grabber. Two FPGA boards are needed as the Virtex-II pro development boards have only 35 valid I/O pins for external interface, while over 50 valid I/O pins are needed for the entire test setup. An external *ADC* is used to digitize the output signal of the image sensor. The *ADC* used is a 12-bit *ADC* from analog devices *(AD9821)* [16]. The frame grabber used in the measurement setup is an *IC-PCI* frame grabber installed in the PC.

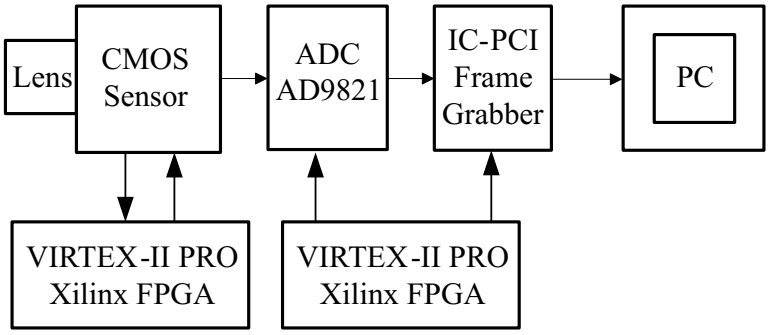


Fig. 3.40 Image sensor characterization setup

DALSA's *LABVIEW* user interface is used to view the acquired image, the row and column waveform, histogram of each frame and average of the frame data. The mean and standard deviation of each row/column or frame is obtained from this interface. The user interface also allows arithmetic operation like averaging and subtraction of pixels or frame data. The black level is an extra offset introduced in the pixel values. The user interface also provides controls to the frame grabber. The frame grabber can be operated in snap or grab mode. When the snap mode is selected, the existing image in the image window will be replaced by a new captured image. In grab mode the images are captured and displayed continuously.

For the experiments, an average measurement of 30 images was obtained to remove the random noise. All measurements were done at room temperature. If a pixel information is represented by $p(x, y)$ where x and y are the row and column number of the sensor array, then the average of a pixel over N frames is given by

$$p_{avg}(x, y) = \frac{1}{N} \sum_{n=1}^{N} p_n(x, y) \tag{3.40}$$

A sensor array of size X by Y pixels is selected for most measurements, where X and Y are the total number of rows and columns respectively. The average of the array of pixels over N frames is then computed as

$$p_{avg}(X,Y) = \frac{1}{XYN} \sum_{n=1}^{N} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} p_n(x,y) \tag{3.41}$$

3.7.2 Measurements

The image sensor has been tested and characterized and shown to be functional. A sample image of a projected text and human hand obtained from the image sensor with an integration time 20ms is shown in the figure 3.41.

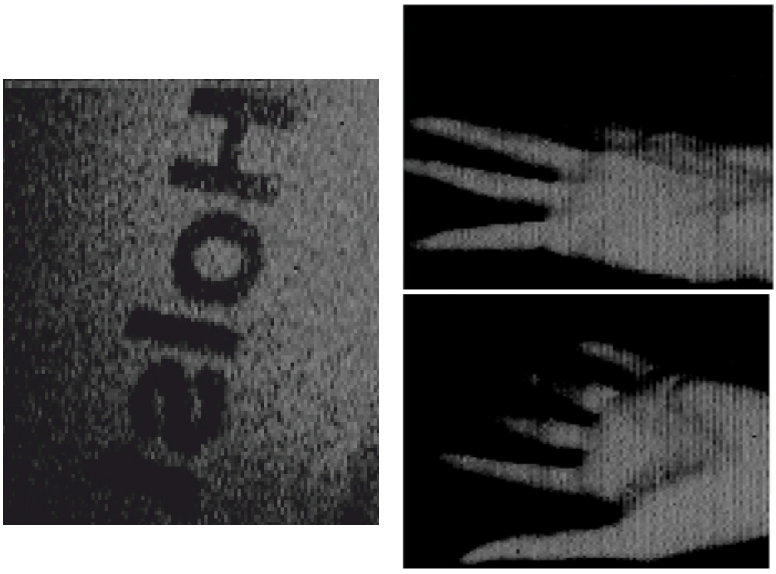


Fig. 3.41 Sample image

3.7.2.1 Photo Conversion Characteristics

The photo-conversion characteristics show the relationship between the output voltage and the incident photon flux. The incident photon flux is a function of the integration time. The photo-conversion characteristic for the image sensor when operated in delta differential sampling mode is shown in figure 3.42. The signal increases linearly before it reaches saturation. The measured maximum output voltage swing from the photo-conversion characteristics is 488mV. This saturation voltage is limited by the column amplifier, which needs to drive the output voltage buffer in figure 3.33. The photosensitivity which refers to the ability to capture the desired details at a given scene illumination is obtained by dividing the slope of the photo conversion characteristics by the gain of the external *ADC* (A_{ADC}). The digitally controlled variable gain amplifier of the *ADC* is set to a digital code of 256 which implies an *ADC* gain (A_{ADC}) of 2.8. The sensitivity is thus calculated to be 13.94mV/ms or 2.14V/lux.s.

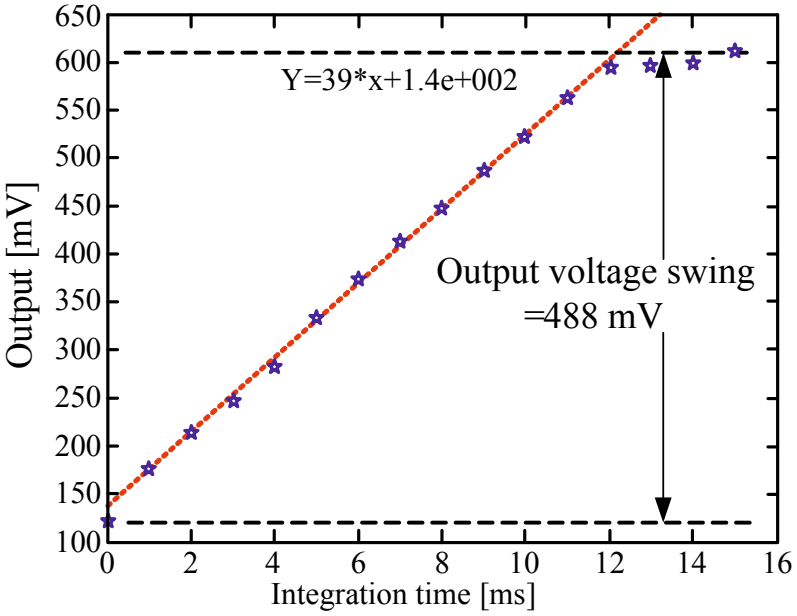


Fig. 3.42 Photo conversion characteristics

3.7.2.2 Photon Transfer Curve *PTC*

The photon transfer curve [10] is a very important method to calculate the various parameters and also to test the overall functionality of the pixel array of the sensor. The Photon Transfer Curve *(PTC)* is a log log plot of noise versus signal to cover the entire dynamic range of the imager. The standard deviation of the noise as a function of the average signal for a group of pixels contained in the array image is plotted.

For a typical array there are three distinct noise regimes. The first regime is the read noise floor measured under dark conditions. Read noise is the noise that is invariant of signal level and thus dominates at very low signal levels. The read noise is the random noise associated with the sensor output amplifier and its signal processing electronics and represents noise sources that are independent of the signal level. The noise floor (read noise) limits the image quality in the dark regions of an image and increases with exposure time due to the pixel dark current shot noise. As the illumination is increased, the dominating noise is the photon shot noise. This is the noise associated with the random arrival of the photons on the array and thus related to the input illumination. The process of random arrival of photons is governed by Poisson statistics, which means that the *rms* value of the shot noise is equal to the square root of the mean number of photons incident on a given pixel. Thus the shot noise profile becomes a straight line with a slope of $1/2$ on the log-log curve. The third regime is associated with fixed pattern or

pixel non-uniformity noise which results from sensitivity differences in each of the pixels. These differences are caused by the manufacturing process and are due to process variations, photo mask misalignments, etc. which cause each individual pixel to have a slightly difference sensitivity to the others pixel on the array. Pixel non-uniformity noise, also known as Photo Response Non-Uniformity or *PRNU*, is directly proportional to the signal level. The measured Photon Transfer Curve *(PTC)* is shown in figure 3.43.

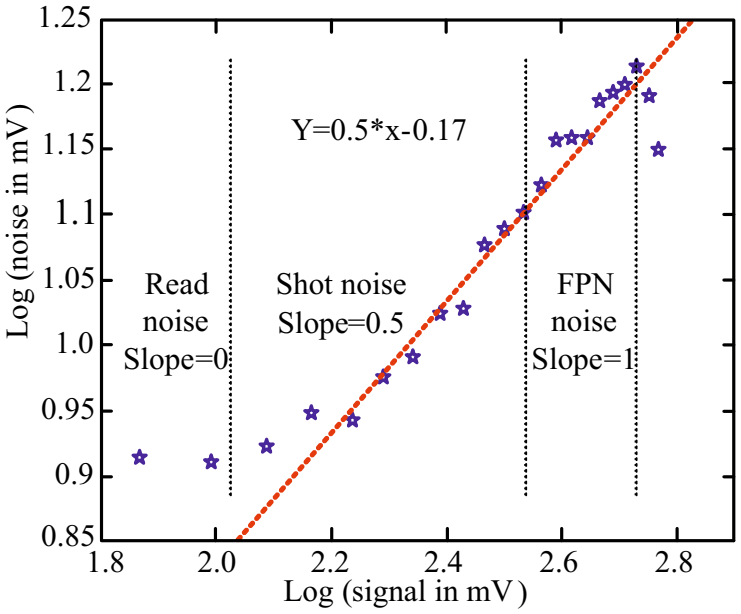


Fig. 3.43 Photon transfer characteristics

The conversion gain *(CG)* of the designed image sensor is measured by extrapolating the photon shot noise back to the X axis intercept (1/100.34 in the case of figure 3.43) and dividing it by the sensor gain (A_{SENSOR} = A_{PIXEL} x A_{OUTPUT} x A_{ADC}). The measured conversion gain is 11.6μV/e^-.The read noise shows the minimum detectable signal which is used for dynamic range calculations. The input referred noise of the pixel is measured to be 18e^-, which is the noise measured at the output of the *ADC* divided by the sensor gain (A_{SENSOR}).

The reset noise is given by

$$\overline{v_{reset}} = \sqrt{\frac{v_T}{CG}} \tag{3.42}$$

where v_{reset} is the reset noise, v_T is the thermal voltage which is 26mV at room temperature of 27°C. The reset noise v_{reset} is calculated to be 47e^-. The reset noise is suppressed by the delta differential sampling operation.

The dynamic range of the sensor is determined by the full well capacity and the readout noise floor. The full well capacity is limited by the voltage swing and the charge storage capacitor. The capacitance at the sense node (C_{FD}) is inversely proportional to *CG* ($C_{FD} = q/CG$), where q is the elementary charge. The calculated sense node capacitance is 14.5fF. The full well or the saturation point of the pixel is the point where the *PTC* curve peaks. The full well from figure 2.34 is derived to be $43186e^-$.

Dynamic range is an important figure of merit for image sensors. The overall dynamic range of the sensor is 45dB. At the sense node of the pixel, the input referred noise is $18e^-$, which gives a dynamic range of 68dB. The dynamic range decreases for increasing integration time due to the effect of dark current.

Thermally generated charge, in addition to the photogenerated charge is accumulated over time at the sense node. This thermally generated current is known as dark current. The dark current charges the pixel capacitance even when there is no light on the pixel, limiting the sensor performance at low illuminations. The amount of dark charge produced is proportional to the integration time. The dark current is measured by shielding the image sensor from any light and acquiring a set of image data for different integration times. The rate of output change or the dark current measured is $18pA/cm^2$ or $0.32mV/ms$ at 27°C.

The total power consumption measured is 100mW at 30fps. The high power consumption is due to the buffer amplifier/comparator inside the pixel which is used for in-pixel image processing. The buffer in the pixel consumes 4.14μW, thus for the complete array the total consumption of all buffers is 67.8mW.

The measured performance parameters of the image sensor are summarized in table 3.3.

Table 3.3 Sensor performance characteristics

Parameters	Measured
Sensitivity	13.94 mV/ms
SNR	33dB
Conversion gain	11.3 μV/e^-
Read Noise (after CDS)	$18e^-$ at pixel level
Dark Current @ 27°C	$18pA/cm^2$
Dynamic Range (after CDS)	68dB at pixel level
Sense Node Cap.	14.5fF
Full well	$43186e^-$

3.8 Summary

- Charge-Coupled Device (CCD) and Complementary Metal-Oxide Semiconductor (CMOS) image sensors are two different technologies used for capturing images digitally.

- The process of photodetection is through photoelectric effect, in which the incident photon energy excites the electrons from the valance band to the conduction band, which are then able to conduct current proportional to the light intensity.
- The p-n junction biased in reverse biased condition is the most popular photodetector. The other variants of the photodiodes are the p^+*/n-well/p-sub*, which has a p^+ pinning layer at the surface.
- The reverse bias helps in forming a depletion region, the photocurrent is integrated over the depletion region capacitance.
- Based on need and performance, the most commonly used pixel types are the passive pixel sensor (PPS), active pixel sensor (APS or 3T pixel), pinned photodiode pixel (4T pixel) and digital pixel sensor (DPS). The PPS is simple in architecture but noisy, APS has less fill factor but low noise compared to PPS and 4T pixel has much lower noise compared to PPS and APS. DPS allows for in-pixel analog-to-digital conversion.
- The logarithmic pixel compresses the light intensity logarithmically and thus help in imaging bright light. The logarithmic pixels have very high dynamic range.
- The performance of a CMOS image sensor is characterized by its quantum efficiency, spectral responsivity, dynamic range, signal-to-noise ratio, conversion gain, noise and dark current.
- The peripheral circuits in CMOS image sensor includes, the addressing shift register or decoders, correlated double sampling (CDS) circuit and analog-to-digital (ADC) converter. The pixel array can be read either in a progressive scam or an interlace scan. The CDS circuit helps in reducing the fixed pattern noise and ADC digitizes the pixel information.
- The ADC can be implemented for chip-level, column-level or pixel-level. In chip-level only one ADC is used for the entire pixel array, in column-level, ADCs are available in each column and in pixel-level the ADC is implemented in pixel like in DPS.
- A CMOS image sensor was designed using the 0.18μm CMOS CIS process from UMC. The pixel is a smart pixel with the ability to perform binarization, store both analog and binary data for current and previous frames. The sensor can be operated in either conventional single charge transfer mode or partial charge transfer mode. The sensor was characterized and was shown to be working.

References

[1] Einstein, A.: Ueber eines die erzeuhung und verwandlung des lichtes betreffendedn heuristischen gesichtspunkt. Annalen der Physik 17, 123–148 (1905)
[2] Sze, S.M.: Physics of Semiconductor Devices. Wiley, New York (1981) ISBN:0471842907
[3] Weckler, G.: Operation of p-n junction photodetectors in a photon flux integrating mode. IEEE Journal of Solid-State Circuits 2, 65–73 (1967)

[4] Gamal, A., Eltoukhy, H.: CMOS image sensors. IEEE Circuits and Device Magazine, 6–20 (2005)
[5] Burkey, B., Chang, W., Littlehale, J., Lee, T., Tredwell, T., Lavine, J., Trabka, E.: The pinned photodiode for an interline transfer CCD imager. In: Proceeding of International Electronics and Device Meeting, Technical Digest, pp. 28–31 (1984)
[6] Teranishi, N., Kohono, A., Ishihara, Y., Oda, E., Arai, K.: No image lag photodiode structure in the interline CCD image sensor. In: Proceeding of International Electron Device Meeting, Technical Digest, pp. 324–3271 (1982)
[7] Decker, S., McGrath, R., Brehmer, K., Sodini, C.: A 256 x 256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output. IEEE Journal of Solid State Circuits 33(12), 2081–2091 (1998)
[8] Liu, X., Gamal, A.E.: Synthesis of high dynamic range motion blur free image from multiple captures. IEEE Transactions on Circuit and Systems–I 50(4), 530–539 (2003)
[9] Shockley, W.: Statistics of recombination of holes and electrons. Physical Review 87(5), 835–842 (1952)
[10] Janesick, J.: Scientific charge coupled devices. SPIE Press (2001) ISBN:9780819436986
[11] Wang, X.: Noise in Sub-Micron CMOS Image Sensors. PhD Dissertation (2008) ISBN:9789081331647
[12] Tian, H., Fowler, B., Gamal, A.E.: Analysis of temporal noise in cmos photodiode active pixel sensor. IEEE Journal of Solid-State Circuits 36(1), 92–101 (2001)
[13] Fossum, E.: CMOS active pixel image sensors. Nuclear Instruments and Methods in Physics Research 395(12), 291–297 (1997)
[14] Baker, R.: CMOS Circuit Design, Layout, and Simulation. IEEE Press, New York (2005) ISBN:047170055X
[15] Xilinx Virtex-II Pro FPGA user guide, http://www.xilinx.com/support/documentation/user_guides/ug012.pdf
[16] AD921:12-Bit 40MSPS ISP, Datasheet Rev0 (2002), http://www.analog.com/static/imported-files/Data_Sheets/AD9821.pdf

Chapter 4
Design of a CMOS Polarization Sensor

Compound eyes found in insects, besides performing their function in forming images and motion detection are also sensitive to other properties of light, i.e. the wavelengths or vector of skylight polarization. Polarization provides additional visual information than intensity and wavelength and also a more general description of light, and therefore it provides richer sets of descriptive physical constraints for the interpretation of the imaged scene.

In this chapter polarization of light waves will be discussed. The mechanism of polarization vision in the compound eyes of insects will also be discussed. Polarization vision is described in section 4.1 and a brief introduction to polarization of light is presented in section 4.2. Polarization vision in compound eyes and polarization cameras are presented in sections 4.3 and 4.4 respectively. The theory behind wire-grid polarizers is discussed in section 4.5 and a one dimensional wire-grid polarized implemented using standard CMOS metal wiring layers is presented in section 4.6. The performance characteristics of such a wire-grid polarizer are also presented in section 4.7. In the final section simulation studies using two dimensional metal layers to selectively transmit specific wavelengths are shown.

4.1 Polarization Vision

The three basic characteristics of light are intensity, wavelength or colour and polarization. The intensity or brightness is an time-averaged expression of the amount of light emitted from the surface of the light source per unit area. The colours of the visible spectrum stretch from violet, with the shortest wavelength, to red, with the longest. Polarization is the plane of vibration of the electric field vector of light.

According to [1], "polarization vision" is the ability to discriminate between two monochromatic lights of the same intensity, but with a different angle of polarization and/or degree of linear polarization. Polarization vision is a generalization of intensity vision and analogous to colour vision. Colour vision is useful in segregating scenes and in recognition of objects while polarization vision discriminates

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 105–155.
DOI: 10.1007/ 978-3-642-34901-0_4 © Springer-Verlag Berlin Heidelberg 2013

reflectances from a scene of similar brightness [1]. The study of polarization is more general than that of intensity or color and can simplify some of the visual tasks (for example region and edge segmentation, material classification etc.) which are more complicated or practically infeasible when limited to intensity and colour information. Because intensity is the compression of polarization information, a polarization camera can also function as a conventional intensity camera. The intensity vision methods can be implemented by such a camera either alone or together with polarization vision methods.

Practically all light that is reflected in most environments is partially polarized, which means that it can be represented by the sum of unpolarized states and completely linear polarized states. Even though the sun itself produces fully unpolarized light, partially polarized light is abundant in natural scenes [2]. The various forms of polarization in natural scenes are discussed in detail in section 3.2. In this book, the term "polarized light" will refer to partially polarized light.

Polarization-dependent optical systems generally fall into three classes:

a) systems which measure polarization properties and are therefore sensitive to any system-induced polarization variation,
b) systems which accurately measure intensities and are therefore sensitive to transmissions which change with polarization state, and
c) systems which route light based on its polarization state and are therefore sensitive to any system-induced polarization variation [3].

A polarization dependent optical system similar to type b, which measures the intensities varying with the polarization state, has been designed in this work.

4.2 Polarization of Light - Basics

Light, viewed classically, can be considered as a transverse wave wherein the underlying oscillations of the electric and magnetic fields happen along directions perpendicular to the direction of propagation of the wave. Longitudinal waves, in contrast, have oscillations confined in the direction of the propagation as in sound waves. Polarization is a phenomenon peculiar to transverse waves. It is the distribution of the electric field in the plane normal to the propagation direction. An unpolarized or randomly polarized electromagnetic wave is one in which the orientation of the electric vector changes randomly.

The mathematical representation of a plane wave propagating in the z direction is

$$\vec{E} = E_0 \cos(kz - \omega t + \varphi_0) \tag{4.1}$$

The electric vector $\vec{E}$ may be decomposed into parallel E_x and perpendicular E_y components as

$$\vec{E} = E_x\vec{x} + E_y\vec{y} \tag{4.2}$$

where E_x and E_y are expressed in the form

$$E_x = E_{x0}\cos(kz - \omega t + \varphi_{x0}) \tag{4.3}$$

$$E_y = E_{y0}\cos(kz - \omega t + \varphi_{y0}) \tag{4.4}$$

From equations (4.3) and (4.4) we have

$$\frac{E_x}{E_{x0}} = \cos(\zeta)\cos(\varphi_{x0}) - \sin(\zeta)\sin(\varphi_{x0}) \tag{4.5}$$

$$\frac{E_y}{E_{y0}} = \cos(\zeta)\cos(\varphi_{y0}) - \sin(\zeta)\sin(\varphi_{y0}) \tag{4.6}$$

where $\zeta = kz - \omega t$. From equations (4.5) and (4.6) we get

$$\left(\frac{E_x}{E_{x0}}\right)^2 + \left(\frac{E_y}{E_{y0}}\right)^2 - 2\left(\frac{E_x}{E_{x0}}\right)\left(\frac{E_y}{E_{y0}}\right)\cos(\delta) = \sin^2(\delta) \tag{4.7}$$

where $\delta = \varphi_{x0} - \varphi_{y0}$ is called the *phase shift*.

Equation (4.7) defines an ellipse, thus the polarized electromagnetic wave of this form is known as *elliptically polarized wave*. The polarization ellipse is very useful since it allows the description of all states of a completely polarized light wave using a single equation. However this representation is inadequate because the light vector propagating in space traces an ellipse or any special form of an ellipse example, the circle or a straight line in a time interval of the order of 10^{-15} sec which is very short to be able to trace the ellipse. Furthermore, the polarization ellipse only represents a completely polarized light and cannot be used to describe either unpolarized light or partially polarized light.

The circular form of the polarization ellipse can be obtained if the phase shift satisfies $\delta = n\pi/2$, where (n=+/-1,+/-3...) and $E_{x0} = E_{y0} = E_0$. Then $\sin(\delta) = \pm 1$ and $\cos(\delta) = 0$, and thus equation (4.7) becomes

$$E_x{}^2 + E_y{}^2 = E_0{}^2 \tag{4.8}$$

Equation (4.8) defines a circle, thus the polarized electromagnetic wave of this form is known as *circularly polarized wave*.

The linear form of the polarization ellipse is obtained if the phase shift $\delta = n\pi$, where (n=0,+/-1,+/-2...) , then $\sin(\delta) = 0$ and $\cos(\delta) = \pm 1$, thus equation (4.7) becomes

$$\left(\frac{E_x}{E_{x0}} + \frac{E_y}{E_{y0}}\right)^2 = 0 \quad or \quad E_y = \pm\frac{E_{y0}}{E_{x0}}E_x \tag{4.9}$$

Equation (4.9) defines a straight line, thus the polarized electromagnetic wave of this form is known as *linearly polarized wave*. A linearly polarized wave has its electric

field vibrating in the same direction at all times at a particular point. A linearly polarized light is often referred to as plane polarized light or simply polarized light.

Linear polarization can be of two types, vertically or *s*-polarized[1] and horizontally or *p*-polarized. A vertically polarized wave is one for which the electric field lies only in the *x*-*z* plane as shown in figure 4.1(a). Horizontally polarized wave is one for which the electric field lies only in the *y*-*z* plane as shown in figure 4.1(b).

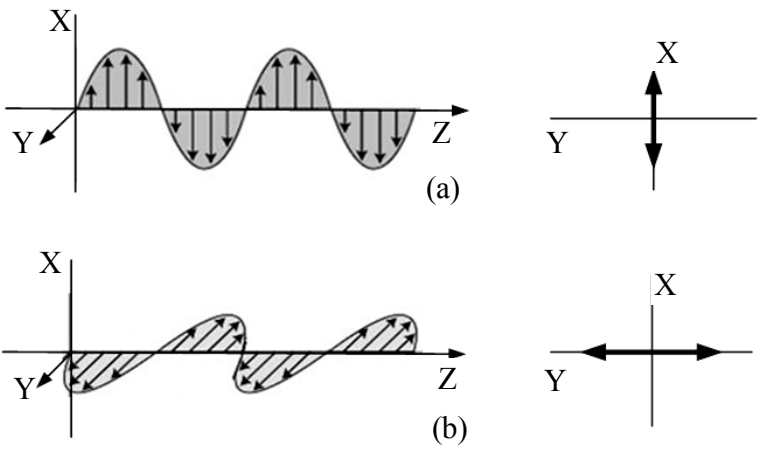


Fig. 4.1 Linear polarization, (a) s-polarized wave (b) p-polarized wave

Elliptically and circularly polarized light is rare in nature [4] and thus the discussion will be limited to the detection and analysis of linear polarization in this book.

4.2.1 Polarization of Light from an Unpolarized Beam

A linearly polarized beam is obtained from an unpolarized beam by removing all waves from the beam except those whose electric field vectors oscillate in a single plane. A polarized light from an unpolarized light can be achieved through a linear polarizer either by the phenomenon of absorption, reflection, refraction or scattering.

4.2.1.1 Polarization by Absorption

The most common technique for producing polarized light is by passing the unpolarized light through a material that absorbs electric field vibrating in all directions except the vibrations in a plane parallel to a certain direction. Such a

[1] The name *s*-polarized comes from German senkrecht meaning perpendicular and *p* is parallel polarized with respect to the plane of incidence.

device is also called a polarizer. One such type of polarizing filter was invented by Edwin H. Land in 1928 and was called Polaroid. A Polaroid consists of long-chain organic molecules that are aligned in one direction, placed in a plastic sheet. They form a closely spaced linear grid, which allows the passage of light vibrating only in the same direction as the grid while absorbing the vibrations in other direction as shown in figure 4.2.

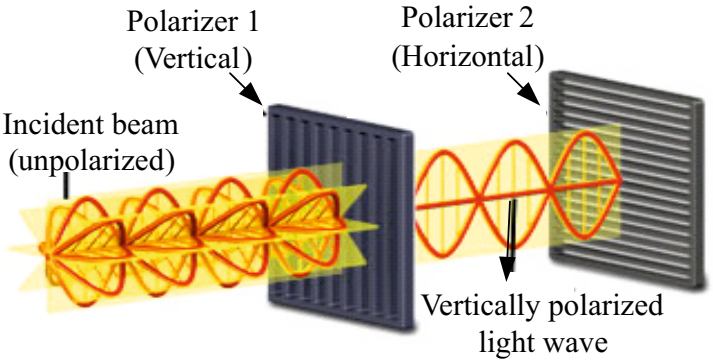


Fig. 4.2 Polarization by absorption [5]

4.2.1.2 Polarization by Reflection

When an unpolarized light beam is reflected from a surface, the reflected light may be completely polarized, partially polarized, or unpolarized, depending on the angle of incidence. For an angle of incidence of 0°, the reflected light is unpolarized, while for other values of angle of incidence, the reflected light is polarized with a vibration direction parallel to the reflecting surface as shown in figure 4.3. If part of the incident light enters the material, the refracted light will also be polarized with a vibration direction perpendicular to the path of the refracted ray.

4.2.1.3 Polarization by Refraction

Optically anisotropic materials have different optical properties in different directions. Double refraction is an optical property where a single ray of unpolarized light entering an anisotropic medium is split into two rays, each travelling in different direction as shown in figure 4.4. Such materials with different indices of refraction associated with different crystallographic directions are known as birefringent materials. Birefringent uni-axial crystalline materials have a unique axis of symmetry, known as the optic axis where no double refraction occurs. The propagation of light along the optic axis would be independent of its polarization; its electric field is everywhere perpendicular to the optic axis and it is called the ordinary- or *O*-ray. The light wave with its electric field parallel to the optic axis is called the extraordinary- or *E*-ray. When an unpolarized light enters a birefringent

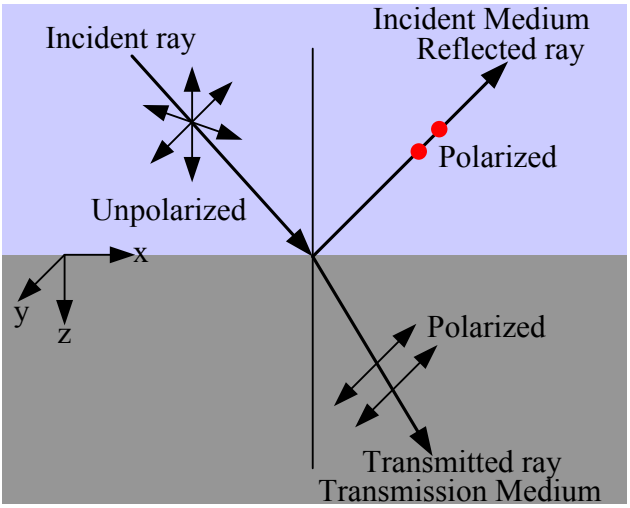


Fig. 4.3 Polarization by reflection

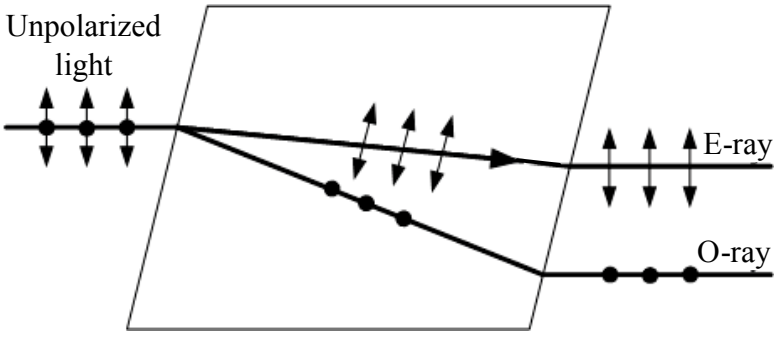


Fig. 4.4 Polarization by double refraction

crystal, the ordinary wave will pass straight through, while the extraordinary wave is displaced by a distance depending on the thickness of the crystal and the angle of refraction.

4.2.1.4 Polarization by Scattering

Polarization or depolarization also occurs when light is scattered while traveling through a medium. When light strikes the atoms of a material, it often sets the electrons of those atoms into vibration. The vibrating electrons then produce their own electromagnetic wave which is radiated outward in all directions. This newly generated wave strikes neighboring atoms, forcing their electrons into

vibrations at the same original frequency. These vibrating electrons produce another electromagnetic wave which is once more radiated outward in all directions. This absorption and re-emission of light waves causes the light to be scattered about the medium. This scattered light is partially polarized.

4.2.2 Polarization of Light - Representation

The state of the polarization as seen from equation (4.7) can be completely defined by four parameters: two amplitudes, the magnitude and the sign of the phase shift. The phase differences are difficult to measure and are not additive when many independent light waves are mixed. For this reason, Stokes parameters introduced in 1852 by Sir George Gabriel Stokes [6] are often used to define the state of the polarization. They provide a phenomenological description of the polarization state of light, involving only the optical intensity that can be experimentally determined. Poincaré sphere was developed by Henri Poincaré in 1892 [7], and is used to visualize the orientation and elongation of the polarization ellipse. In this section, the Stokes parameters are related to intensity and the orientation angles of the Poincaré sphere.

The polarization ellipse discussed in section 4.2 describes the polarized light in terms of its amplitude. The amplitude of the optical field cannot be observed, but it is possible to observe and measure its intensity, which is equal to the time average of the square of the field amplitudes. This can be done by taking a time average of the time dependent quantities in equation (4.7). The time average of the field components are defined by the following equation:

$$\langle E_x(z,t)E_y(z,t)\rangle = \lim_{T\to\infty}\left(\frac{1}{T}\int_0^{\infty} E_x(z,t)E_y(z,t)dt\right) \tag{4.10}$$

where T is the time of the measurement. To evaluate equation (4.7) the denominator is first removed by multiplying by the factor $4{E_{x0}}^2{E_{y0}}^2$. We then have

$$\begin{aligned}&4{E_{y0}}^2\langle {E_x}^2(z,t)\rangle + 4{E_{x0}}^2\langle {E_y}^2(z,t)\rangle \\ &\qquad - 8E_{x0}E_{y0}\langle E_x(z,t)E_y(z,t)\rangle\cos(\delta) = (2E_{x0}E_{y0}\sin\delta)^2\end{aligned} \tag{4.11}$$

Using equations (4.3), (4.4) and (4.10), the time averages in equation (4.11) are found to be

$$\begin{aligned}\langle {E_x}^2(z,t)\rangle &= \frac{1}{2}{E_{x0}}^2 \\ \langle {E_y}^2(z,t)\rangle &= \frac{1}{2}{E_{y0}}^2 \\ \langle {E_x}^2(z,t){E_y}^2(z,t)\rangle &= \frac{1}{2}E_{x0}E_{y0}\cos\delta\end{aligned} \tag{4.12}$$

Substituting equation (4.12) in (4.11) and rearranging we get

$$\left({E_{x0}}^2 + {E_{y0}}^2\right)^2 - \left({E_{x0}}^2 - {E_{y0}}^2\right)^2 - \left(2E_{x0}E_{y0}\cos\delta\right)^2 = \left(2E_{x0}E_{y0}\sin\delta\right)^2 \quad (4.13)$$

From equation (4.13) we can define

$${S_0}^2 = {S_1}^2 + {S_2}^2 + {S_3}^2 \quad (4.14)$$

where

$$\begin{aligned} S_0 &= {E_{x0}}^2 + {E_{y0}}^2 \\ S_1 &= {E_{x0}}^2 - {E_{y0}}^2 \\ S_2 &= 2E_{x0}E_{y0}\cos\delta \\ S_3 &= 2E_{x0}E_{y0}\sin\delta \end{aligned} \quad (4.15)$$

The parameters S_0, S_1, S_2 and S_3 are known as the Stokes polarization parameters for a plane wave.

The first parameter expresses the total intensity of the optical field while the remaining three parameters describe the polarization state of light. The Stokes vector helps in the measurement of

i.) the intensity of light
ii.) the degree of linear polarization with respect to vertical and horizontal axes
iii.) the degree of linear polarization with respect to the axes oriented at +45°and -45°
iv.) the degree of left and right circular polarizations

The Stokes parameters are obtained by measuring the transmitted intensity through a set of linear polarization filters. The Stokes parameters can be measured using a retarder and a linear polarizer. The retarder (with retardance ϕ) advances the electric field in the *x*-direction by $\phi/2$ and retards the electric field in the *y*-direction by $\phi/2$.

$$\begin{aligned} E_x &= E_{x0}e^{\phi/2} \\ E_y &= E_{y0}e^{-\phi/2} \end{aligned} \quad (4.16)$$

The transmitted electric field with the polarization transmission axis at θ is given by

$$E_x = E_{x0}e^{\phi/2}cos\theta + E_{y0}e^{-\phi/2}sin\theta \quad (4.17)$$

The intensity *I* can be obtained from EE^*.

$$I(\theta,\phi) = \frac{1}{2}[\big((E_x E_x^* + E_y E_y^*)\big) + (E_x E_x^* - E_y E_y^*)cos2\theta + (E_x E_x^* + E_y E_y^*)cos\phi sin2\theta + i(E_x E_x^* - E_y E_y^*)sin\theta sin2\phi]$$

$$= \frac{1}{2}[S_0 + S_1 cos2\theta + S_2 cos\phi sin2\theta + S_3 sin\theta sin2\phi] \tag{4.18}$$

where S_0, S_1, S_2 and S_3 are expressed in terms of the complex amplitudes. With no retarder ($\phi = 0°$), the Stokes parameters S_0, S_1 and S_2 can be easily determined by varying the transmission axis of the linear polarizer by $\theta = 0°, 45°$ and $90°$. S_3 can be determined with retarder having $\phi = 90°$ and the polarization transmission axis of $\theta = 45°$. From equation (4.18) we can obtain

$$\begin{aligned} I(0°,0°) &= \frac{1}{2}[S_0 + S_1] \\ I(45°,0°) &= \frac{1}{2}[S_0 + S_2] \\ I(45°,90°) &= \frac{1}{2}[S_0 - S_1] \\ I(45°,90°) &= \frac{1}{2}[S_0 + S_3] \end{aligned} \tag{4.19}$$

From the above equations the Stokes parameters can be solved as

$$\begin{aligned} S_0 &= I(0°,0°) + I(90°,0°) \\ S_1 &= I(0°,0°) - I(90°,0°) \\ S_2 &= 2I(45°,0°) - I(0°,0°) - I(90°,0°) \\ S_3 &= 2I(45°,90°) - I(0°,0°) - I(90°,0°) \end{aligned} \tag{4.20}$$

One of the main advantages of the Stokes formalism is that it allows for the determination of the degree of polarization *(DOP)* of any optical signal directly from its Stokes components. In most cases, an electromagnetic wave consists of the incoherent superposition of a fully polarized wave and of a fully unpolarized wave. The *DOP* quantifies that fraction of the total optical signal that is actually polarized, with:

$$DOP = \frac{\sqrt{{S_1}^2 + {S_2}^2 + {S_3}^2}}{S_0} \tag{4.21}$$

The *DOP* of a fully polarized optical signal is equal to one, whereas the *DOP* of an unpolarized optical signal is zero. In a linearly polarized light beam,the circular and elliptical polarizations are zero. Its degree of polarization is thus often referred to

as degree of linear polarization *(DOLP)*. The degree of linear polarization of a light beam is defined by:

$$DOLP = \frac{S_1}{S_0} \tag{4.22}$$

To visualize the polarization ellipse of a propagating polarized light wave characterized by an azimuthal angle and an ellipticity angle, Poincaré sphere as shown in figure 4.5 is used. On the sphere, the longitude lines represent the ellipticity angle χ. and the latitude lines represent the azimuthal angle ψ. The Stokes parameters are the Cartesian coordinates of a space in which any completely polarized light beam is represented by a point on a sphere with unit radius around the origin.

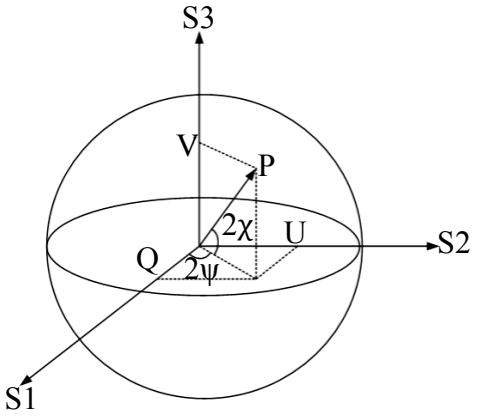


Fig. 4.5 Poincaré sphere representation

The Cartesian coordinates *(x,y,z)* of the Stokes parameters are mapped to the spherical coordinates (I, ψ, χ) of the Poincaré sphere [8] of unit radius using the following equations

$$\begin{aligned} S_1 &= \delta \cos(2\psi) \cos(2\chi) \\ S_2 &= \delta \sin(2\psi) \cos(2\chi) \\ S_3 &= \delta \sin(2\chi) \end{aligned} \tag{4.23}$$

The Stokes parameter S_0 represents the radius of the Poincaré sphere, while S_1, S_2 and S_3 represent the coordinates of a point on the surface of the sphere. Each point on the surface of the sphere corresponds to a state of the polarization of the light. Right circular polarization is represented at the north pole and left circular polarization at the south pole of the sphere. Linear polarization is represented along the equator of the sphere. Elliptical polarization of all possible azimutal and ellipticity angles are represented by other points on the surface of the sphere.

The polarization azimuth angle ψ and the ellipticity angle χ can then be derived from equation (4.23) as:

$$\sin(2\chi) = \frac{S_3}{\sqrt{{S_1}^2 + {S_2}^2 + {S_3}^2}}$$
$$\tan(2\chi) = \frac{S_2}{S_1} \tag{4.24}$$

The parameter χ varies from +45° to -45°. It is positive for right-handed polarization forms and negative for left-handed polarization forms. The parameter ψ varies from 0° to 180°, it is 0° for parallel polarization and 90° for perpendicular polarization.

4.3 Polarization Cameras

The polarization state of light can be easily visualized using polarization filters. A polarization filter has the property to either transmit light vibrating in one direction producing linearly polarized light or light with a sense of rotation producing circularly polarized light. Basically a polarization filter transforms polarization information into intensity difference information. This is important because humans are able to perceive intensity information but not polarization information.

A polarization camera is a generalization of the conventional intensity camera. It extends the capability of conventional imaging by providing polarization information about a scene, besides the intensity information. The spatial variation in intensity or polarization can be used to obtain high contrast images. Polarization cameras improve the overall contrast of the image even in low light environment, wherein the polarization contrast compensates the low intensity contrast. The significant improvement in the image contrast has increased the attention of polarization imaging in various applications.

Polarization imaging cameras has been developed for visible, mid infra-red and long infra-red wavelength regions. In polarization imaging to obtain parameters describing the polarization state of the monochromatic light, a set of measurements has to be taken from different positions of the polarizers, waveplates and photodetectors. Therefore, data acquisition time depends on how fast the polarizers, waveplates and photodetectors are moved to their various measurement positions. In addition, the precision of positioning the polarizers, waveplates and photodetector influences the results of measurements.

State of the art polarization image sensors consist of a photodetector array, such as a CMOS or a CCD sensor array, and a single or multiaxis micropolarizer array to measure the polarization information in real time. Based on the position of the micropolarizers and the mode of data acquisition, the current state-of-the-art polarization image sensors can be divided into two types: time domain multiplexed mode sensors and spatial domain multiplexed mode sensors.

The first group of polarization sensors includes standard imaging sensors coupled with electrically or mechanically controlled polarization filters [9]. It is operated in

time domain multiplexed mode, where in, the externally coupled linear polarizer is rotated in front of either a single photodetector or a complete photodetector array. The main disadvantages these systems is that a mechanical or electrical rotation of the polarizer filter is required, thus it is not suitable for real time design. Multiple images with different orientation of the polarization filter are required for the measurement of partial linear polarization, increasing the frame time by the total number of orientation measurements desired. These cameras are not suitable if the camera is in motion or the scene being imaged is continuously changing. Further as the precision of the position of the polarizers effects the measurement results, these systems often produces erroneous results. However these systems have higher spatial resolution as the whole imaging array images only one orientation at one time.

To overcome the disadvantages of the externally coupled micropolarizer and to realize real-time polarimetric imaging, liquid crystal displays that are operated at high frequencies have been proposed. The modulation of the liquid crystal display allows capturing light to sequentially sample the Stokes parameters in more or less simultaneous manner. However, one of the serious drawbacks with using liquid-crystal displays for polarization imaging is the lack of an inherent calibration [10]. The other idea proposed in the literature is to split the incoming light into multiple parts, each of these light part is then incident on an imaging sensor coupled with polarization filters of varying transmission axis. This is similar to the use of colour splitting prism to separate three primary colours for colour imaging in classical imaging tubes [11] and solid state imaging sensors [12]. The disadvantage of this method is the necessity of multiple detectors to detect the polarization information.

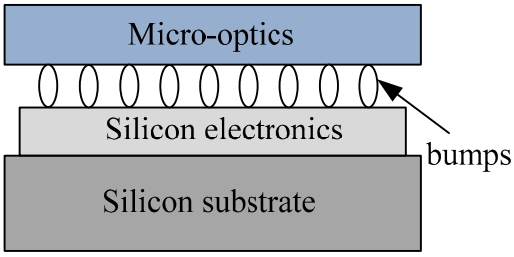


Fig. 4.6 Micro-polarizer integrated over silicon substrate

The second group of polarization sensors includes image sensors where the linear micro-polarizer is either fabricated directly on the sensor array or on a substrate and then flip-chip bonded with the sensor array. A layout for typical flip-chip bonded polarization detection sensor is shown in figure 4.6. The polarization filters are located or fabricated on the micro-optic layer. The micro-optic filter is then positioned over the imaging array. Each polarization filter in the micro-optic layer spatially connects to an underlying pixel and thus each pixel receives polarization information of only one direction. The polarization filter on the micro-optic layer is

shown in figure 4.7. The 2 x 2 array of the polarization filter usually corresponds to three polarization filters and a circular polarization filter. The division of available filters into three or four different linear polarization filters leads to reduction of the image resolution, however this structure makes it possible to detect polarization in real-time. Monolithic integration of micropolarization filters with CMOS imaging sensors is currently a subject of intense research. A plane coated with micrometer-scale polarizing elements in a variety of orientations is called a micropolarizer array. They operate in spatial domain multiplexed mode where, the linear polarizers are oriented in different directions with respect to the photodiodes and are directly fabricated on top of the photodiode. Micropolarizers have been fabricated with metal grids and organic materials. Incorporating pixel-pitch-matched polarization filters at the focal plane for visible spectrum has been explored with birefringent materials [13] and thin-film polarizers [14], [15].

Photolithographically patterning polymer polarization film to make micropolarizer arrays was proposed by Faris [16]. Polymer film micropolarizer arrays have been successfully deployed in display applications, but their use in imaging is more challenging. For imaging, the thickness of the polarizing element has to be much smaller than the pixel size to have a large light collection angle. Polymer polarizers are made by mechanically stretching the polymer film. It is difficult to manufacture uniform stretched films with a thickness of less than 10μm [17]. Another main disadvantage of the thin-film polymer polarization imaging sensors has been the

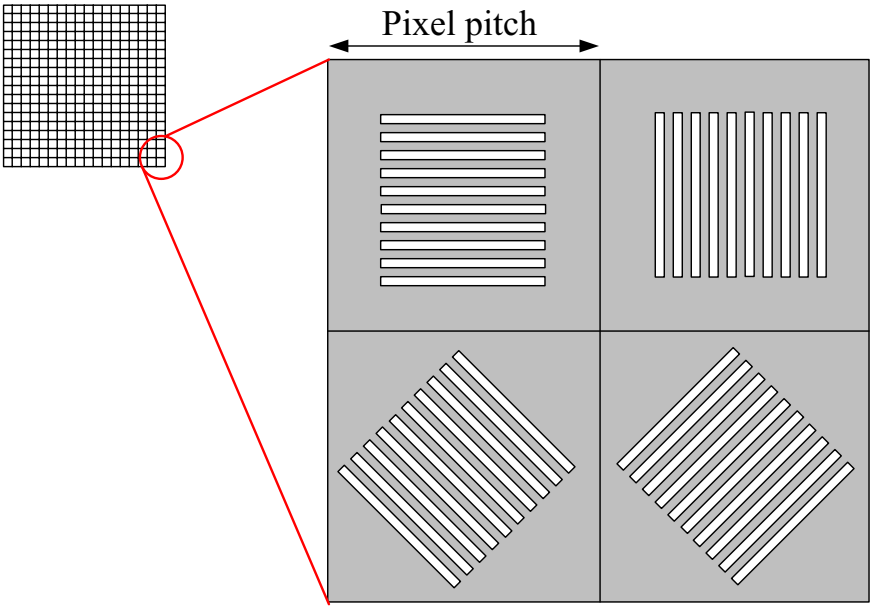


Fig. 4.7 2x2 array of polarization filters in the micro-optical layer

complex fabrication techniques required to merge CMOS technology with polymer filters. These fabrication steps are prone to misalignment errors, which can degrade the sensing capabilities of the polarization imaging sensor. Furthermore, the thick multilayer filter array, which is around 10 micron of thickness, is prone to optical crosstalk and limits the extinction ratios of the sensors [18]. The spatial orientation of the linear polarizer filter results in reduced spatial resolution. However since the linear polarizers are directly incorporated on top of the photodiode, the temporal resolution is increased allowing for faster polarization imaging and also contributes to the compactness of the design. The precision of the system also increases as the position of the polarizer is fixed.

Metal-grid polarizers for electromagnetic radiation have been used for a long time and can be dated back to Hertz [19]. Metal-grid micropolarizer arrays are often used for infrared imaging polarimetry [20], but they were not very effective in the visible wavelength range. However with the scaling of the modern CMOS processes the metal planes normally used for interconnections when specifically aligned are now becoming suitable for polarization transmission [21], [22]. Wire-grid polarizers are discussed in detail in section 3.5.

Table 4.1 summarizes the discussion on different types of linear micropolarizers.

Table 4.1 Comparison of linear polarizers

	External filter	Organic micro-polarizer	Metallic wire grid polarizer
Operating mode	Time domain	Spatial domain	Spatial domain
Mechanical intervention	Required	Not required	Not required
Extinction ratio	Very high	High	Low
Spatial resolution	High	Low	Low
Temporal resolution	Low	High	High
Additional fabrication required?	No (externally coupled)	Yes	No (Standard CMOS metal layers are used)

4.4 Wire-grid Polarizer

A wire-grid is used to denote a planar structure composed of a series of parallel wires or strips made of a conducting material. They are metallic gratings with periods much smaller than wavelength of the incident electromagnetic wave. Wire-grid polarizers are commonly used polarizing elements for the wavelength range of 1 to 1000μm.

Wire-grid polarizers are inorganic optical polarization devices made of aluminum. Aluminum is usually chosen for the wire-grid polarizer in the visible region of the electromagnetic spectrum as it is highly reflective and absorbs only a small

portion of the incident electromagnetic wave in the visible region. Furthermore they are able to withstand high temperatures, while the performance of organic polarizers is typically degraded at high temperatures [23]. The aluminum wire-grid polarizer also provides high transmission efficiency and polarization extinction ratio.

Normally, such a structure would function as a simple diffraction grating, but when the pitch or period of the wires is less than half the wavelength of the incoming light, it becomes a polarizer. Wire-grid polarizers are extensively used as polarization sensitive elements in various applications because of their excellent polarization performance and planar structure that allows them to be easily pixilated and integrated into other optoelectronics devices. A schematic view of a wire-grid polarizer is shown in figure 4.8.

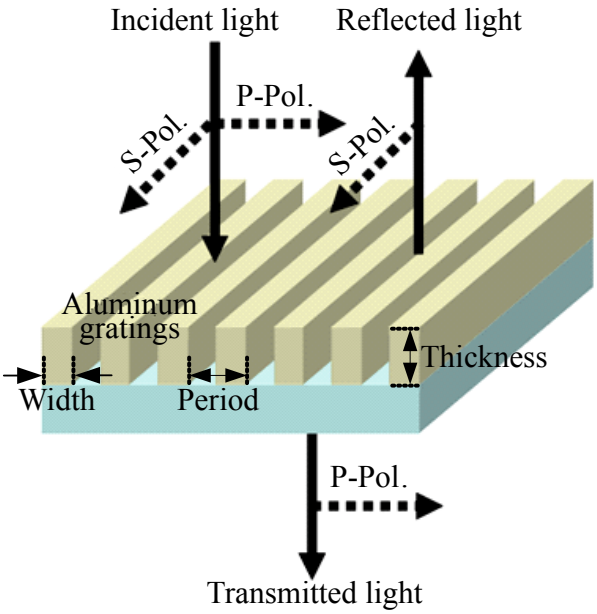


Fig. 4.8 Wire-grid polarizer [24]

The polarization effects that occur when an electromagnetic wave is transmitted through a wire-grid having a grid pitch smaller than the wavelength of the incident electromagnetic wave was first observed by Heinrich Hertz in 1888 [19]. When light is incident on the wire-grid polarizer, the light which passes through the grating structure is affected depending on the size, shape and spacing of the grating structure. If the electric field vector, in a linearly polarized light is parallel to the wire-grid (*s*-polarization), it excites electron motion in the wire-grid. This causes the electromagnetic wave to be reflected back from the surface of the metal while some energy is lost because of joule heating in the wires. If the electric field is perpendicular to the wire-grid (*p*-polarization), the electron motion is confined

within the wire, as the force on the electrons is applied in a perpendicular direction. Thus the electromagnetic wave is transmitted with little or no loss in energy. The wire-grid polarizer transmits electromagnetic waves whose electric field is perpendicular to the grid wires (*TM* polarization) and reflects or radiates off the electromagnetic waves with an electric field parallel to the wires (*TE* polarization). The wire-grid polarizer acts like a mirror for reflected *s*-plane polarized light and as a dielectric for the *p*-plane polarized light as shown in the figure 4.8. For the one dimensional grating shown in figure 4.8, appropriate selection of dimensions of the wire-grid will result in behavior as a polarizer, waveplate or polarization dependent filter.

The theoretical explanation for the behavior observed for the transmission of the electromagnetic wave through a wire-grid polarizer was provided by Horace Lamb in 1898 [25]. For a grating of strips, the transmission coefficients were shown to be

$$T_{\parallel} = \frac{y^2}{1+y^2} \;\; ; \;\; y = \frac{2D}{\lambda} ln \left[\sec \pi \left(\frac{1}{2} - \frac{A}{D} \right) \right] \tag{4.25}$$

$$T_{\perp} = \frac{1}{1+x^2} \;\; ; \;\; x = \frac{2D}{\lambda} ln \left[\sec \left(\frac{\pi A}{D} \right) \right] \tag{4.26}$$

where A is the half-width of the strip, D is the grating space, λ is the wavelength of the incident light, $T_{\parallel}$ and $T_{\perp}$ are the transmission coefficients of the electric field parallel and perpendicular to the grid wires respectively. The equation shows the dependence of the transmitted intensity on the size of the wire-grid. The equations also show that for $\lambda \gg D$, the transmission coefficient of the electric field parallel to the wire-grid is nearly zero; thus the transmitted electric field is completely polarized with only the orthogonal electric field present.

If the wire-grid shown in figure 4.8 is considered to be a diffraction grating and is illuminated by a light of wavelength λ, it generates diffraction orders according to the grating equation [26]

$$sin\varphi_m = \frac{1}{n} \left(sin\varphi_{in} + m\frac{\lambda}{p} \right) \tag{4.27}$$

where p is the grating period, n is the refractive index of the substrate material, φ_{in} and φ_m are the angle of incidence and the diffraction of the m-th order, respectively. Geogre Bird and Maxfield Parrish first suggested that if the pitch of the wire-grid polarizer was small enough it can be used as a polarizer in the visible region. The wire-grid polarizer structure acts as a diffraction grating whose interaction with incident light can be modeled by rigorous application of the diffraction grating theory and the boundary conditions of Maxwell's equations. For the wire-grid, polarizing effect only in the 0th order will be used, and thus a grating period where

only the 0th diffraction order propagates (a zero order grating) at any angle of incidence $0° < \varphi_{in} < 90°$ is necessary. This condition can be described by

$$p \leq \frac{\lambda}{2n} \tag{4.28}$$

Further if *t* is the grating width and *a* is the grating period *(t+D)*, the refractive indices of the *TE* wave n_{te} (the electric vector parallel to the grating grooves) and the *TM* wave n_{tm} (the electric field vector normal to the grating groove) are expressed as:

$${n_{te}}^2 = R{n_1}^2 + (1-R){n_2}^2 \tag{4.29}$$

$${n_{tm}}^2 = \frac{{n_1}^2{n_2}^2}{R{n_2}^2 + (1-R){n_1}^2} \tag{4.30}$$

where n_1 is the dielectric constant of the grating material, n_2 is the dielectric constant of the fill material and *R* is the ratio of the grating width to the grating period *R=t/a*.

The key parameters that determine the performance of a wire-grid polarizer are its period (pitch), linewidth, and depth of the grid elements. The geometrical parameters of the wire-grid polarizer are shown in figure 4.9.

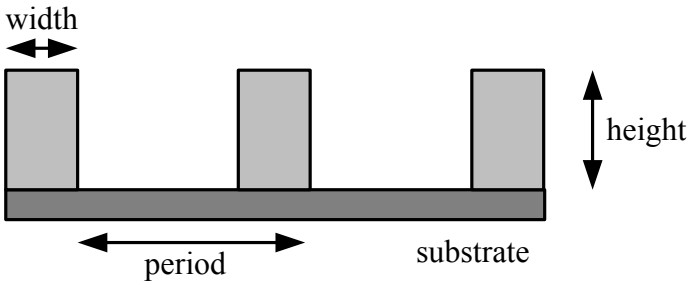


Fig. 4.9 Schematic illustration of the cross section of a wire-grid polarizer with its main geometrical parameters

The depth of the grid elements (thickness of aluminum) should be enough to be optically opaque and it should also be able to achieve the desired transmission and extinction ratio. The wire height of a visible spectrum wire-grid polarizer is typically between 100nm and 200nm. The wire-grid polarizer aspect ratio is the ratio of the wire thickness to the width of the wire. For a visible spectrum wire-grid polarizer the aspect ratios usually range from 2:1 to 25:1. The linewidth is the width of the polarizer material. Typically the linewidth for visible spectrum wire-grid polarizers is approximately 60nm, although this may vary based upon the desired performance of the polarizer. The duty cycle is defined as the ratio of line width and the pitch.

It is basically the amount of space one wire occupies compared to the pitch. The duty cycle range generally used to manufacture visible spectrum wire-grid polarizers is from 40% to 60%. The wire-grid polarizer duty cycle affects the reflection and transmission efficiencies as well as the reflection and transmission extinction ratios.

Pitch (period) is defined as the distance between the center of two consecutive polarizer wires. The wavelength for which a given grid will act as a useful polarizer is $\lambda = 2d$, where d is the grid spacing [27]. The pitch needs to be at least three times smaller than the smallest wavelength that is being polarized for optimum performance. For example, to efficiently polarize a spectrum with a minimum wavelength of 400nm, the wire-grid polarizer should have a pitch of at most 133nm.

4.4.1 Fabrication of Wire-grid Polarizer

The commercially available wire-grid for infra-red region consists of thin film metallic grid patterned on infrared detecting substrates like ZnSe, Ge, BaFe, Polyster and Polyethylene. The two most commonly used method to pattern the metallic wire-grid are the interference lithography and contact printing.

Interference lithography is used for fabricating periodic and quasi-periodic patterns over large areas. Interference lithography is a simple process where two coherent beams interfere to produce standing wave, which is recorded in a photoresist. The spatial grating period as low as half the interfering light can be fabricated using interference lithography technology. A wire-grid fabricated using interference lithography is shown in the figure 4.10.

Ultraviolet (UV) lithography, such as contact lithography is used for grid patterns with larger periods (for example with periods larger than 4μm). In contact lithography, the metallic nano-pattern from a mold which is directly transferred

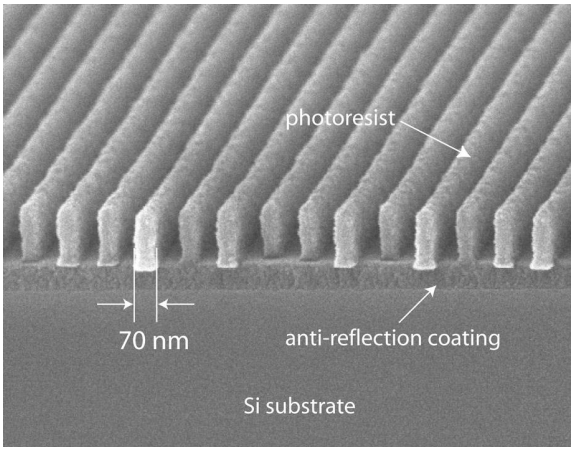


Fig. 4.10 Wire-grid fabricated using interference lithography with a grating width of 70nm with a grating pitch of 170nm [28]

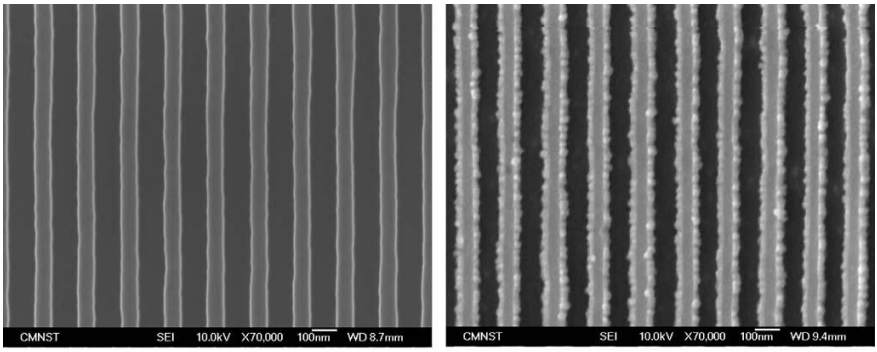


Fig. 4.11 SEM images of (a) a Si mold wire-grid with a linewidth of 60nm and a pitch of 170nm and (b) transferred metallic patterns [29]

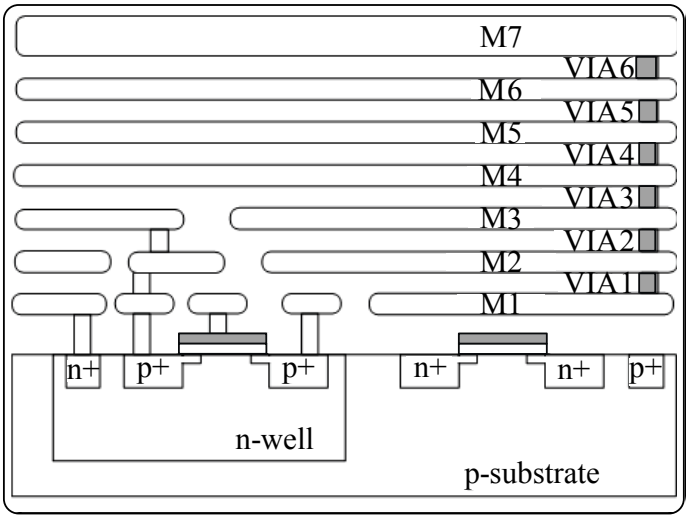


Fig. 4.12 Standard CMOS metal layer arrangement

to a polymer substrate. This process is advantageous as it transforms the metallic pattern to a polymer substrate in one step. A wire-grid fabricated using a Si mold and contact lithography is shown in the figure 4.11. However, these fabrication steps are complicated and indirect. In most CMOS manufacturing processes, the electrical connections between pixel components are routed through a series of metal layers that are between the imaging element of the system (lens) and the photodetector. The cross-sectional view of the metal layers of the standard CMOS process technology from 0.18μm UMC is shown in the figure 4.12.

It is preferred to have the opening between the light sensitive portion of the photodetector and the lens to be free of any metal interconnects to increase the quantum efficiency. In modern manufacturing processes, the metal lines within the pixel are comparable or smaller than the wavelength of visible light (400nm to 700nm). At these scales, the metal layers are not as opaque to visible light as bulk metal, and variations in the pattern control the probability that incident light at a particular wavelength will influence the photodetector. With the scaling of CMOS technologies, the distance between metal layers also scales, opening up the possibility of using them in a grid structure for the absorption of electromagnetic waves.

4.4.2 *Transmittance Efficiency and Extinction Ratio of Wire-grid Polarizers*

When randomly polarized light is transmitted through an ideal wire-grid polarizer, the electric fields orthogonal to the wires will be transmitted and electric fields parallel to the wire will be reflected, as seen in section 4.4. In practice however, wire-grid polarizers transmit some undesired parallel electric fields while reflecting some desired orthogonal electric field and also absorbing some of the incident light. Thus the transmitted electric field through a wire-grid polarizer can be modeled as having two components, an electric vector associated with the parallel electric field and an electric vector associated with the orthogonal electric field that are each scaled by a scalar value. The scalar value is proportional to the square root of the intensity associated with each polarization. Equation (4.31) is the resulting electric vector equation for unpolarized light passing through a wire-grid polarizer.

$$\vec{E} = t\left(\hat{p}.\vec{E}\right)\hat{p} + c\left(\hat{r}.\vec{E}\right)\hat{r} \tag{4.31}$$

where E is the resultant electric field after transmission through the wire-grid polarizer and the scalar value for the parallel electric field is t and the scalar value for the orthogonal electric field is c.

The irradiance of the plane polarized light passing through a polarizer shown in figure 4.13 is given by the "law of Malus" [30].

If the transmission axis of the polarization filter is parallel to the x-axis of the incident polarized light the transmitted intensity is given by equation (4.32).

$$I = I_0 \cos^2(\varphi) \tag{4.32}$$

where I is the transmitted intensity, I_0 is the maximum transmitted intensity and φ is the angle between the polarizer's transmission axis and the plane of polarization of the incident light. Thus the maximum intensity that can be sensed by the detector is dependent of the angle of the transmission axis of the polarization filter and would result in a cosine wave with varying transmission axis angle.

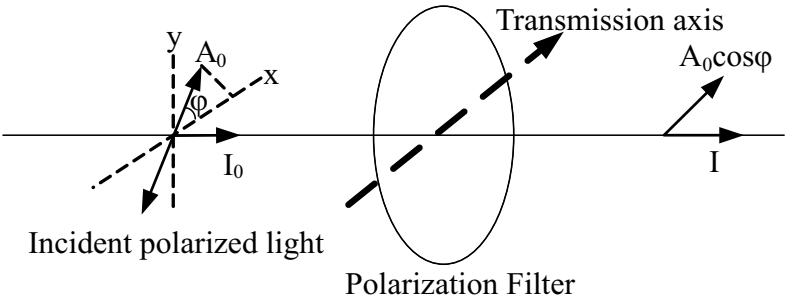


Fig. 4.13 Intensity transmission through a polarizer, Malus law

Based on the transmission of intensity at a given transmission axis of a linear polarizer, linear polarizers are characterized by two main specifications: transmittance efficiency and extinction or contrast ratio.

The wire-grid polarizer's *transmittance efficiency* is the fraction of the total incident light that is transmitted through the linear polarizer. The maximum transmission in the linear polarizer is obtained when the polarizer orientation is aligned parallel to the orientation of the linear polarized component and the minimum is obtained when the polarizer orientation is aligned perpendicular to the linear polarized component of light.

The transmission of the *p* and *s* polarized wave through a wire-grid depends on the duty cycle of the wire-grid as seen in section 4.4. Figure 4.14 shows the FDTD (finite-difference time-domain) simulations of *s* and *p* polarization transmission and reflection through a 500nm period wire-grid made of Aluminum for a 1.55μm wavelength. The figure shows that as the duty cycle increases the *p* and *s* transmission decreases [31].

The transmission efficiency increases with decrease in the grating period as shown in figure 4.15. While as the duty cycle is increased, the wires get wider and closer (distance between them decreases), it attenuates both the *p*-polarization and the *s*-polarization and thus the transmission efficiency decreases as shown in figure 4.16 [32].

The effect on transmission due to the change in metal grating height is shown in the figure 4.17. The TM transmission for different wavelength decreases as the metal height increases [32].

Figure 4.18, shows the performance of the wire-grid with varying groove depth and angle of incidence for an 550nm incident electromagnetic wave. The depths of the wires are varied from 10nm to 400nm while the duty cycle is fixed to 45% with a period of 144nm and width of 65nm of the wires [33]. When the plane of the incident beam is perpendicular to the wire-grid, the *p*-polarization transmission is the highest, because the electric field is perpendicular to the wire-grids. When the plane of incident beam is not perpendicular to the wire-grid, the electric field is not perpendicular to the wire-grid and thus *p*-polarization transmission drops. In fact when the plane of incident is not perpendicular, *p*-polarized light has a component

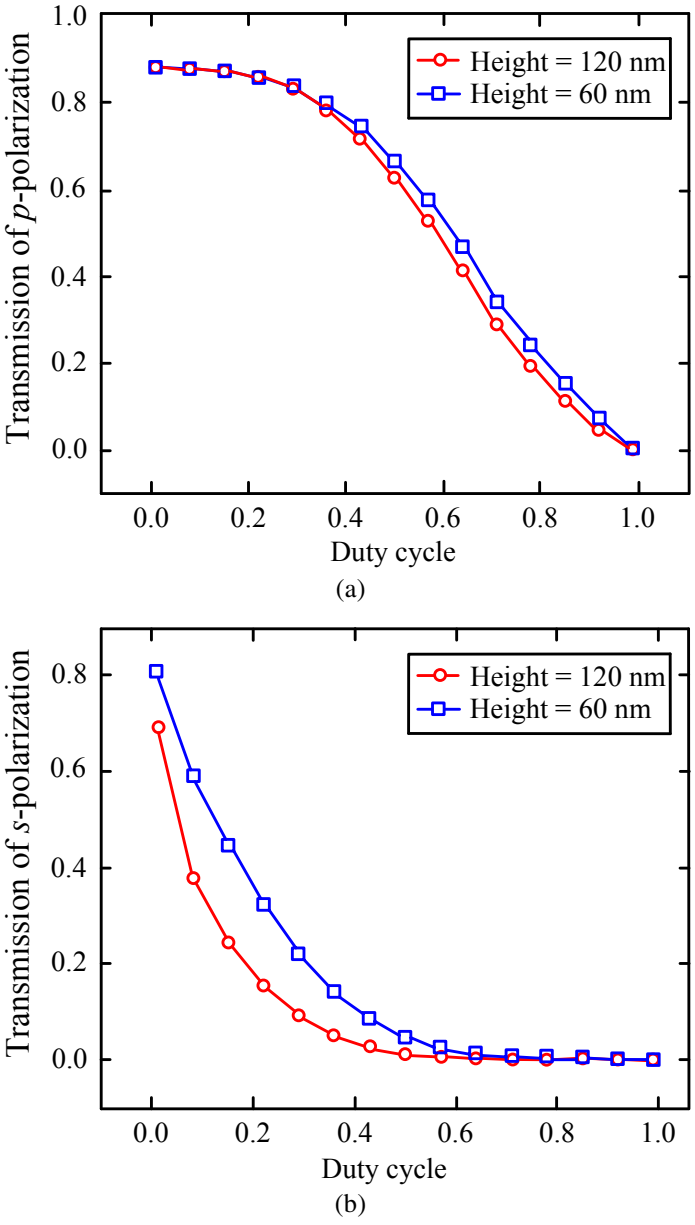


Fig. 4.14 Simulated transmission through a wire-grid polarizer as a function of duty cycle for various aluminum metal heights(a) p-polarization, (b) s-polarization [31]

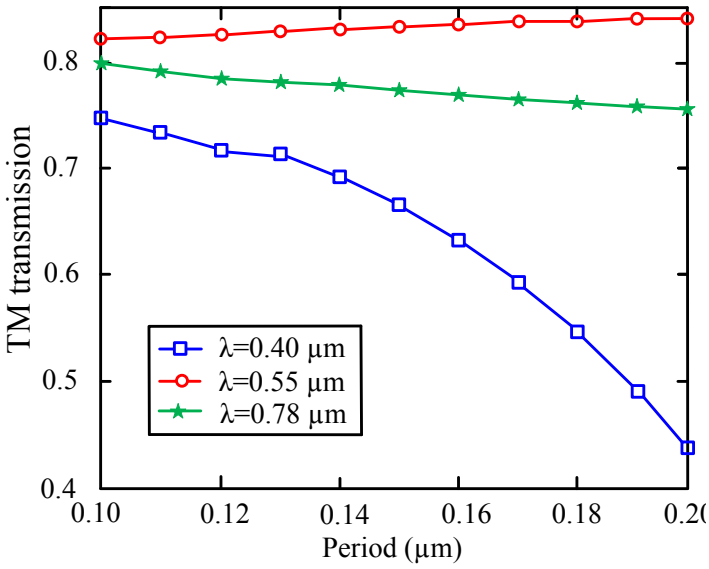


Fig. 4.15 TM transmission as a function of grid period [32]

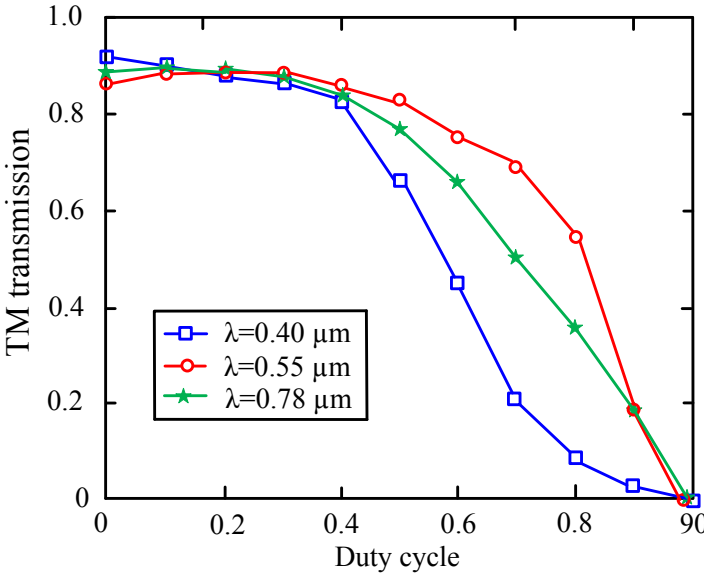


Fig. 4.16 TM transmission as a function of duty cycle [32]

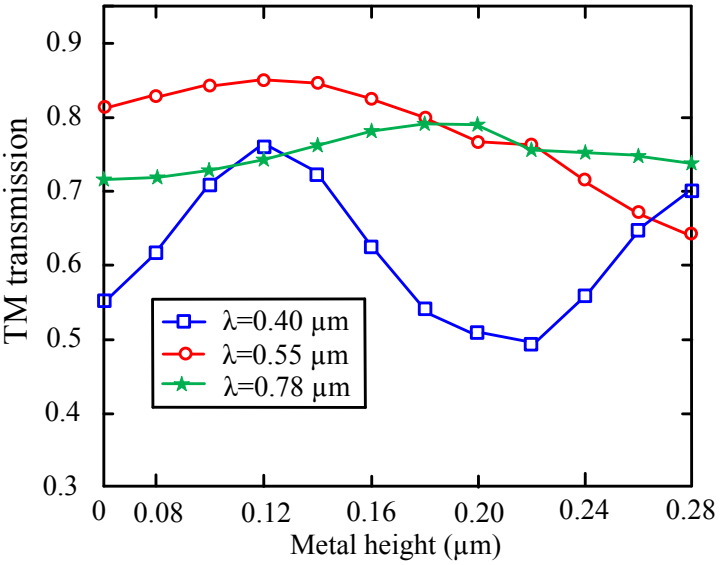


Fig. 4.17 TM transmission as a function of metal height

parallel to the wire-grid, which increases the reflection of the *p*-polarized light. The simulated transmission of TM polarized light with varying angle of incidence is shown in the figure 4.19. The transmission profile decreases with increase in angle of incidence of the electromagnetic wave on the wire-grid polarizer.

Besides the geometrical parameters, the grating material also affects the properties of the wire-grid polarizer, as materials significantly differ in their optical properties. Figure 4.20 shows the theoretical simulation of the transmission of *p*-polarization for the grating made of aluminum, gold and silver. The figure 4.20 shows that aluminum as a grating material shows the best transmission performance in the whole spectral range. As mentioned earlier, aluminum is thus preferred choice for fabricating wire-grid polarizer because it allows maximum *p*-polarization transmission and reflects most of the *s*-polarization.

The transmission profile is affected by the refractive index of the material. The refractive index of aluminum, silver and gold as a function of wavelength is shown in the figure 4.21, 4.22 and 4.23. The figures shows that the real and imaginary part of the refractive index of the aluminum increases linearly with wavelength in the visible region.

The transmittance of the transmagnetic (*TM*) polarization component is also important, since poor *TM* transmittance would require a highly sensitive sensor element even with an excellent *ER*. In an ideal wire-grid polarizer 50% of the light intensity would be reflected in the *s*-polarization state and 50% transmitted in the *p*-polarization state. The polarizer would have a transmission efficiency of 100% and an infinite extinction ratio. Since unpolarized natural light consists of

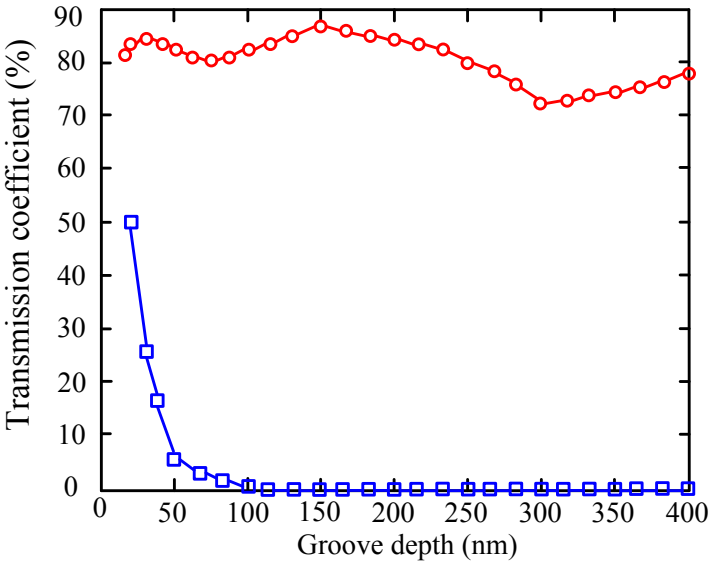


Fig. 4.18 Reflection and transmission of p and s polarization for varying groove depth for 550nm wavelength at normal angle of incidence [33]

two orthogonal linear polarizations, polarization of light can only provide 50% of the light in the desired polarization state.

The *extinction ratio (ER)* is a measure of the polarization contrast of a linear polarizer and it is defined as the intensity ratio between the maximum and minimum throughputs of the polarizer when the wire-grid is aligned and crossed with respect to the light beam. The accuracy of the wire-grid polarizer's extinction ratio measurement depends on several factors such as the detector precision, analyzer alignment, stray light variations and light source stability. In an ideal testing environment with no light source or detector drift or background noise variations, the extinction ratio can be simply defined as

$$ER = \frac{I_p}{I_s} \tag{4.33}$$

where I_p is the maximum transmittance parallel to the polarized beam and I_s is the minimum transmittance perpendicular to the plane of the polarized beam. The contrast ratio for polarizers varies depending on the polarizer, ranging from around 10 to over 10^7.

Extinction ratios can be further enhanced using two wire-grid polarizers in series with parallel transmission axis, where the overall extinction ratio is the product of the extinction ratios of the individual polarizers. The extinction ratio is also a

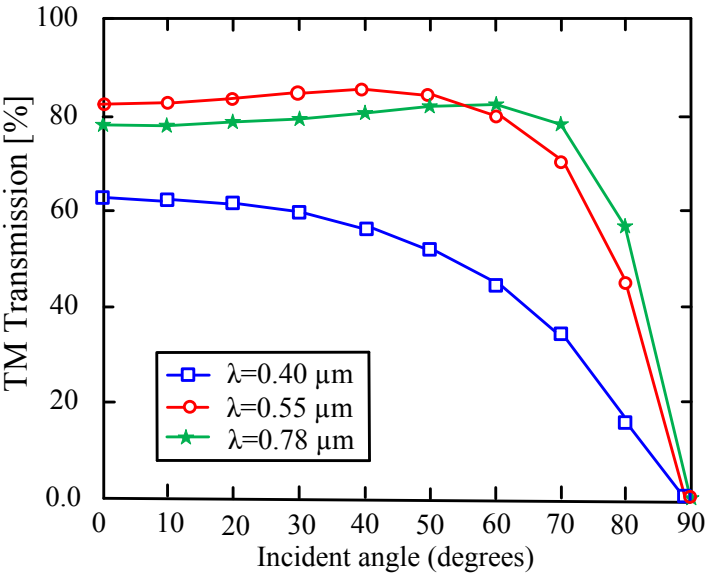


Fig. 4.19 TM transmission for various angle of incidence of the electromagnetic wave on the wire-grid polarizer [32]

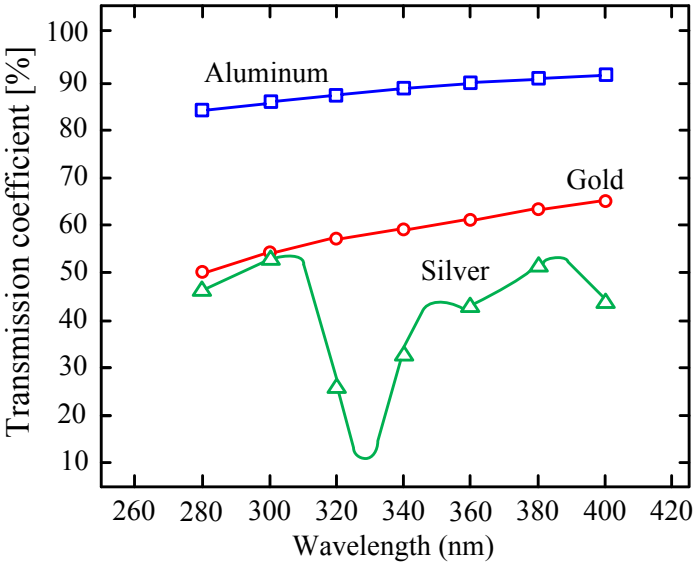


Fig. 4.20 TM-Transmission for a grating structure with various grating materials for a grating period of 100 nm, a metal width of 50nm and a height of 150nm [26]

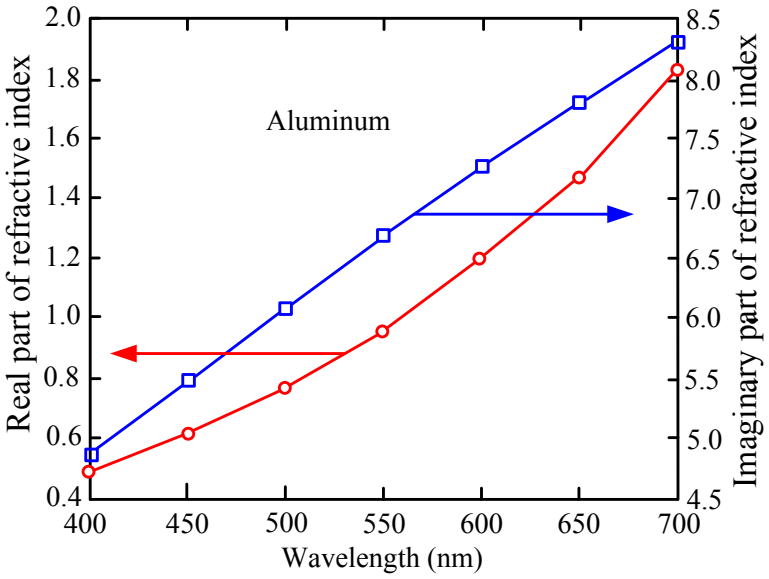


Fig. 4.21 Real and imaginary part of the refractive index of aluminum [33]

function of groove depth, the fill factor of the metal, the period of the metal grating and also the shape of the wires in the wire-grid.

Figure 4.24 shows the simulated polarization extinction ratio as a function of the duty cycle and metal height of the wire-grid made of aluminum metal for an electromagnetic wave of wavelength 1.55μm. Figure 4.25 shows the variation of the extinction ratio with varying duty cycle and wavelength of the electromagnetic wave. The extinction ratio increases with increasing duty cycle.

Figure 4.26 shows the simulated polarization extinction ratio for varying shapes of the grating in the wire-grid. The anisotrpically etched grating have a period of 500nm, the metal thickness is 100nm with a duty cycle of 0.5, while the trapezoidal edged model has a period of 500nm, the bottom width is 250nm and the top width is 60nm [31]. The figure shows that the extinction ratio decreases for trapezoidal shapes as the top width is lower.

Figure 4.27 shows the simulation results for the variations in polarization extinction ratio as the pitch of the aluminum grating in a wire-grid polarizer is changed. In figure 4.27, the height and the duty cycle of the aluminum grating are set to 140nm and 0.5, respectively. It is clearly visible from the figure that for a higher polarization extinction ratio, the pitch of the wire-grid has to be very small.

The variation in the extinction ratio, with changing metal height is shown in the figure 4.28. The extinction ratio increases with increase in the metal height. This can be explained as follows: the *s*-polarization sees the wire-grid as an absorbing film, increasing the metal height increases the absorption. The *p*-polarization sees the

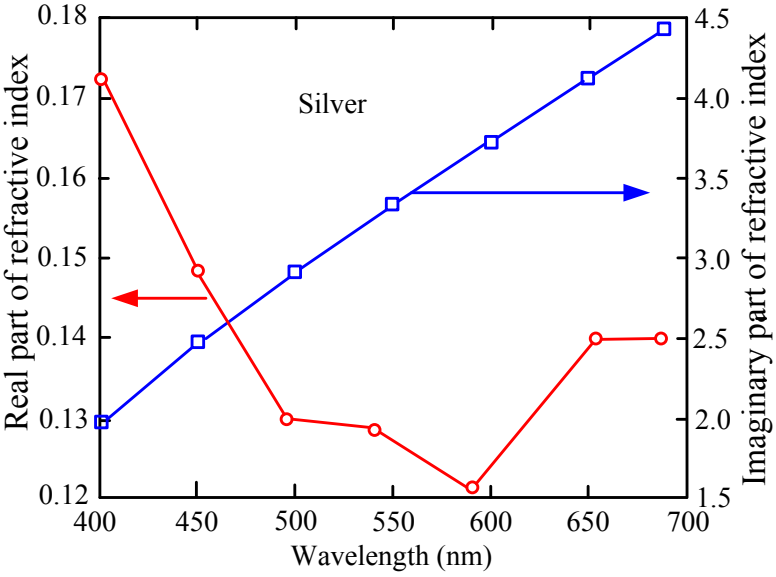


Fig. 4.22 Real and imaginary part of the refractive index of silver [33]

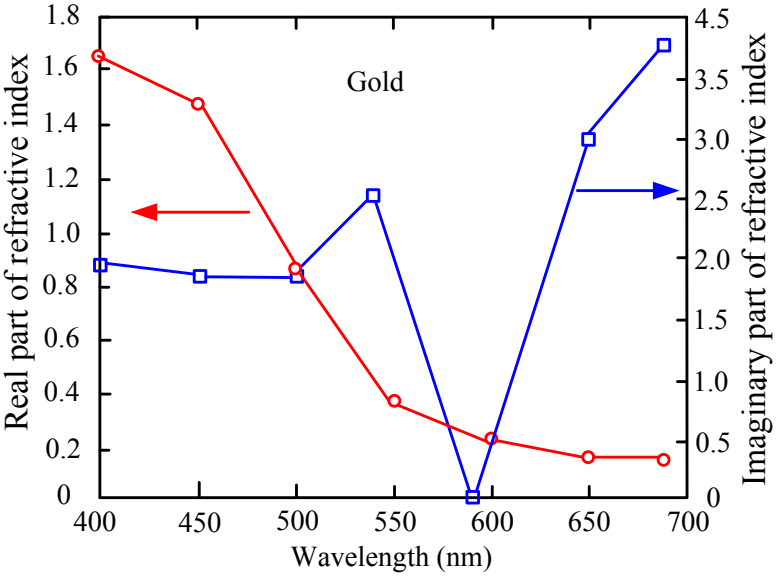


Fig. 4.23 Real and imaginary part of the refractive index of gold [33]

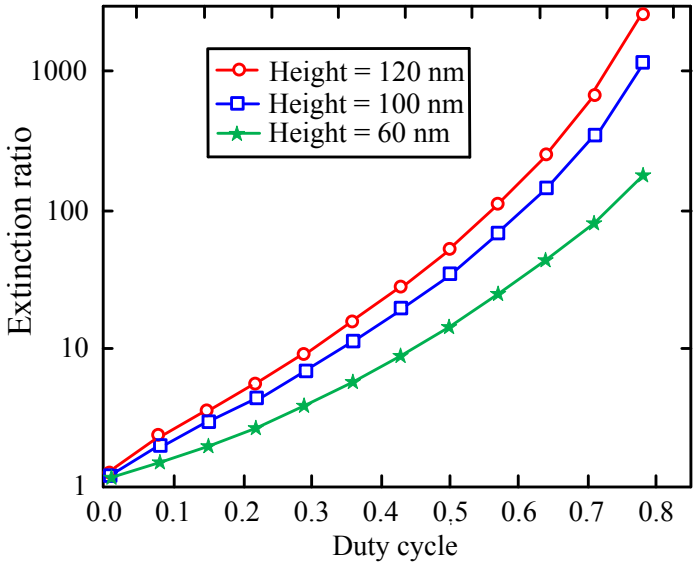


Fig. 4.24 Simulated polarization extinction ratio as a function of duty cycle for various aluminum wire heights for a wire-grid with a 500nm period [31]

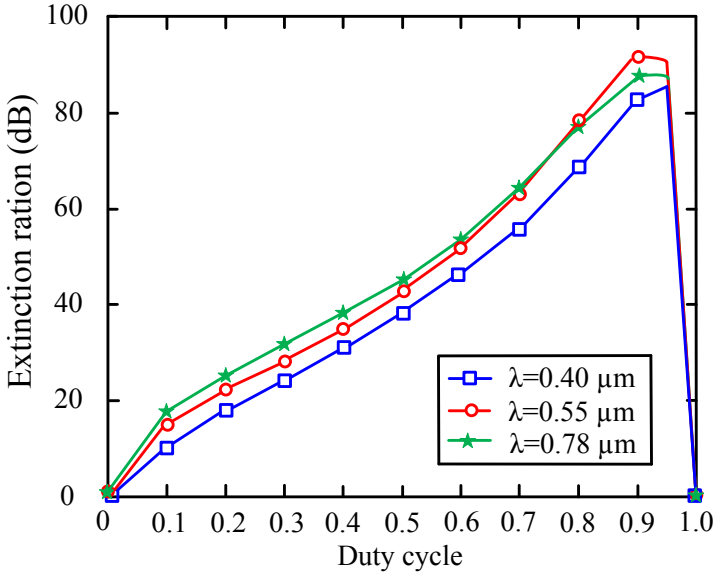


Fig. 4.25 Simulated polarization extinction ratio as a function of duty cycle and wavelength of the electromagnetic wave [32]

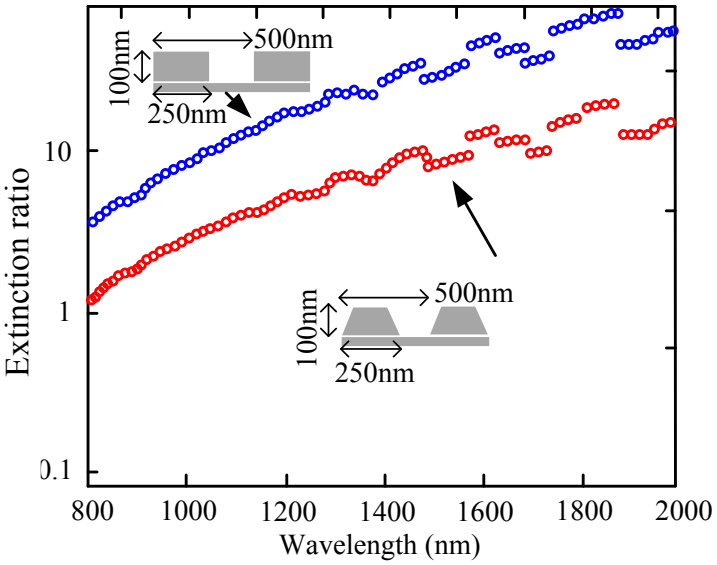


Fig. 4.26 Simulated polarization extinction ratio for aluminum wire-grid with a rectangular and trapezoidal shapes device for a wire-grid with 500nm period [31]

grating as weekly absorbing dielectric material. These results in increased extinction ratio with increase in the metal height.

The variation of the extinction ratio for wire-grid fabricated with different materials is shown in figure 4.29. It is observed that for Aluminum, highest extinction ratio is obtained and thus is the preferred choice os metal for fabrication of metallic wire-grids. From the figure it is clear that the aluminum wire-grid has a better extinction ration than the wire-grid made of gold and silver. In general, metals with refractive indices where the imaginary part is large and the real part is small gives better extinction ratio and thus performs better as wire-grid.

The transmission extinction ratio has a weak angular dependence. However the reflection extinction ratio has a very strong angular dependence. At Brewster angle, the wire-grid polarizer will have the highest reflection extinction ratio. The Brewster angle for a wire-grid is a function of the orientation of the wire-grid, and this can be varied between 0°and 90°by varying the height and width of the wire-grid. The variation of extinction ratio with the variation in the incidence angle of the electromagnetic wave is shown in figure 4.30.

A novel wire-grid polarization image sensor combining CCD imaging technology with nanotechnology has been recently been presented [34]. The aluminum nanowire optical filters are monolithically integrated directly on top of the imaging array as close as possible to the photosensitive element using a modified procedure of interference lithography. Fabricating these metallic nanowires with subwavelength dimensions has become feasible with the maturity of nanofabrication

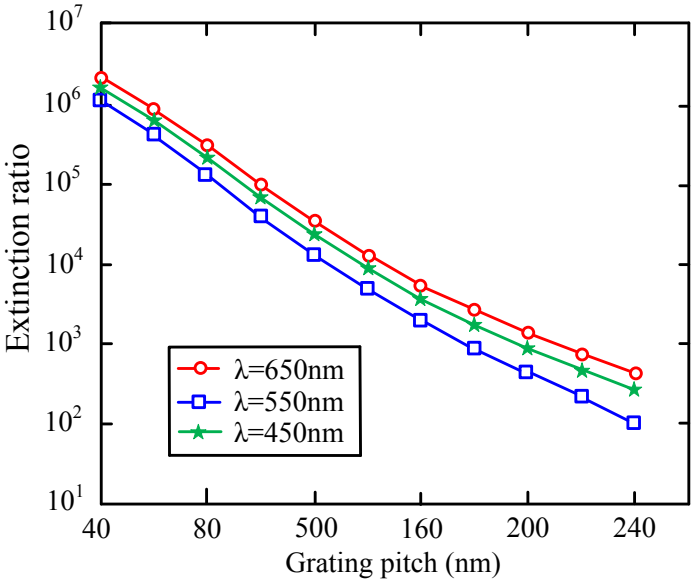


Fig. 4.27 Extinction ratio variation with pitch of the aluminum wire-grid polarizers [24]

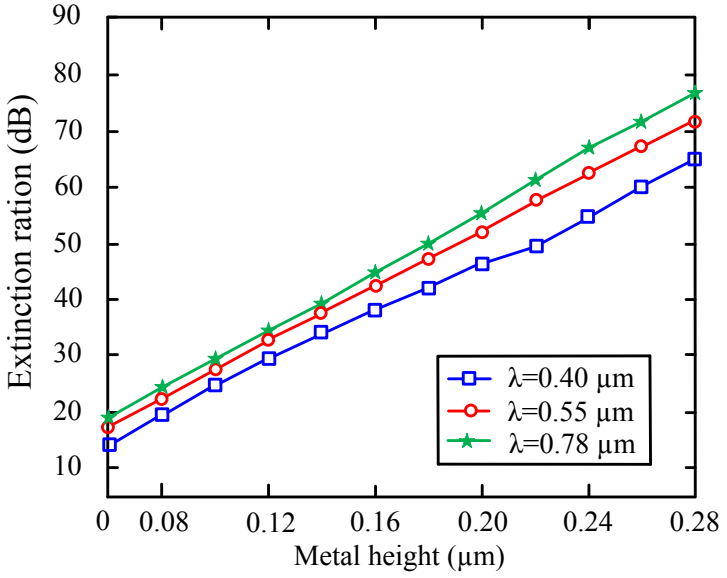


Fig. 4.28 Extinction ratio variation with height of the aluminum wire-grid polarizers [32]

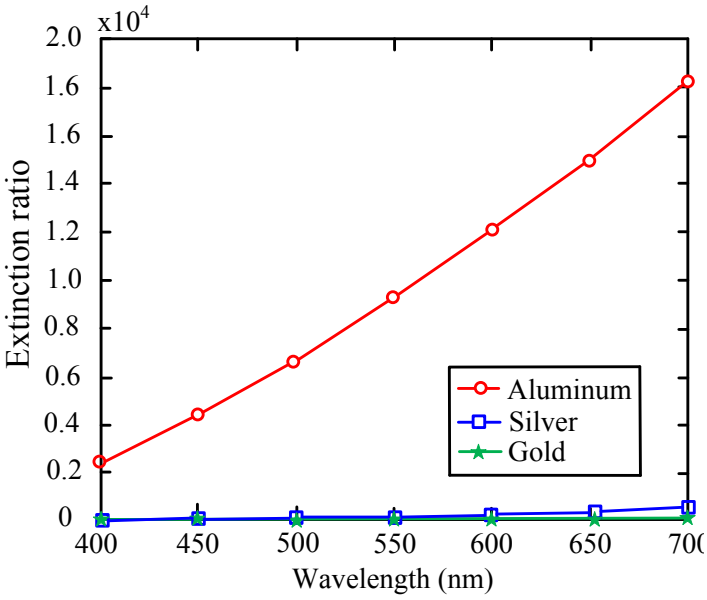


Fig. 4.29 Extinction ratio variation with wavelength for various metal wire-grid polarizers [33]

technologies and techniques [23]. The optical nanowire polarization filter made by first depositing a 70nm thin film of aluminum followed by 30nm thin film deposition of SiO_2 using *e*-beam evaporation. The maximum extinction ratio reported for an incident light wavelength of 625 $\pm$ 5nm and is 58 $\pm$ 1.2 across the imaging array. The extinction ratios for green (515 $\pm$ 5nm) and blue (460 $\pm$ 5nm) wavelengths are 44 $\pm$ 0.7 and 30 $\pm$ 0.5 respectively. For incident light (625 $\pm$ 5nm wavelength) perpendicular to the surface of the imaging sensor, the extinction ratios are maximum and equal to 58 [34].

A high resolution liquid crystal micropolarimeter for visible linear (0°, 90°) and circular polarization (right handed) imaging in real time has also been reported [35]. An extinction ratio as high as $\sim$1100 at the wavelength of 500nm is reported. The polarization filter is implemented by micro-patterning a liquid crystal layer on top of a 45° oriented ultra-thin metal wire-grid polarizer. Large area aluminum nanowire grid with 40nm/78nm space and 118nm pitch, fabricated using full wafer immersion interference lithography is also available [23]. The wire-grid is shown to perform as a highly efficient optical polarizer for deep ultraviolet down to $\sim$250nm. The average extinction ration at 300nm is reported to be $\sim$400.

The metallic wire-grid linear polarizers can also be made using the metal planes normally used for interconnections in modern CMOS processes with gate-lengths of 180nm or below. Because the size of these metal planes are smaller than the wavelength of visible light, they show a unique transmission behavior for optical waves. By proper design of the structures with gratings or holes in one or more metal layers, optical functionality such as polarization transmission can be achieved.

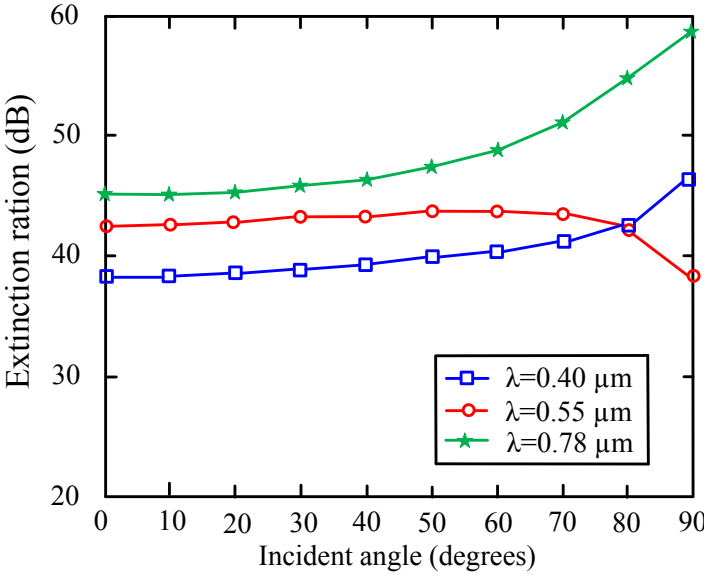


Fig. 4.30 Extinction ratio variation with angle of incidence of the electromagnetic wave on the wire-grid polarizers [32]

A polarization-analyzing CMOS image sensor using 0.35μm 2-poly 4-metal standard CMOS technology was reported by [21], [22]. The sensor featured a monolithically embedded polarizer. The angle of the embedded polarizer was varied to realize a real time absolute measurement of the incident polarization angle. The extinction ratio of the embedded polarizer reported was as small as 2. Although the pixel-level performance was limited, the estimation schemes based on the variation of the polarizer angle provided promising performance for uniform illuminations.

Besides transmittance efficiency and extinction ratio, linear polarizers are also characterized using the degree of polarization *(DOP)*. The degree of polarization of a wire-grid polarizer in terms of maximum and minimum transmittance is given by

$$ER = \frac{I_p - I_s}{I_p + I_s} \tag{4.34}$$

where I_p is the maximum transmittance parallel to the polarized beam and I_s is the minimum transmittance perpendicular to the plane of the polarized beam.

Typically, for a totally polarized light a *DOP* close to one is obtained while for a completely unpolarized light a *DOP* of near zero is desired.

4.5 Design of a Polarization Sensor

In section 3.5, its shown that to selectively absorb the electromagnetic waves in the visible spectrum the grid pitch should be very small. Thus a lower technology nodes,

much lower than 0.18μm, would be ideal to design the wire-grid micropolarizers. However the CIS (CMOS Image sensor) process of UMC with 3-metal layers provide only 0.18μm photodiodes, the photodiodes for technology nodes below 0.18μm are not very well characterized. In a 0.18μm standard CMOS process the minimum allowable distance between two metals is 0.24μm. A one-dimensional wire-grid polarizer implemented using the bottom metal layer (M1) is shown in the figure 4.31.

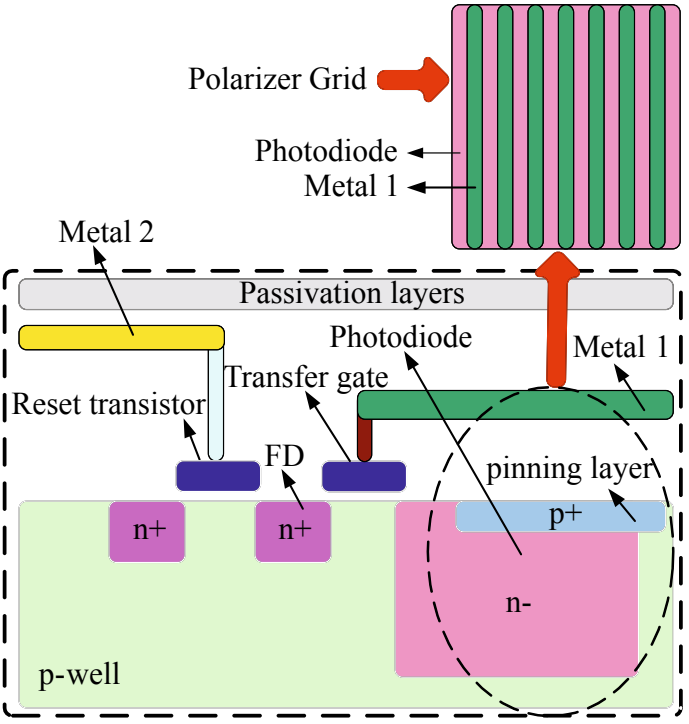


Fig. 4.31 Pixel cross-section and top view of the implemented wire-grid polarizer

The wire-grid polarizer is implemented over the pixel described in chapter 3. The thickness of the metal layer M1 used to create the wire-grid polarizer is 0.35μm while the width of metal wire is 0.24μm. The choice of the thickness and the width of the metal wire are limited by the choice of the technology. The aspect ratio for the designed wire-grid micro-polarizer is 1.5:1. The duty cycle of the designed wire-grid micro-polarizer is 50% with a pitch of 0.48μm. The wire-grid parameters are summarized in table 4.2. The grid metal thickness is fixed by the CMOS process metal thickness. Based on the selected grid width and height, the incident *p*-polarized light whose electric field is parallel to the grating is mostly transmitted and the *s*-polarized whose electric field is perpendicular to the grating is mostly transmitted and detected by the photodiode.

Table 4.2 Wire-grid parameters

Metal thickness	0.35μm
Metal width	0.24μm
Aspect ratio	1.5:1
Grid pitch	0.48μm
Duty cycle	50%

Figure 4.32(a) shows the implementation of the polarization analyzing CMOS image sensor operated in spatial mode. The wire-grid micro-polarizer was created with varying orientations over the pixel array to enable real time measurement of the polarization information. The active pixel array of 128 by 128 pixels was split into two regions of 64x128 pixels. One region of 64x128 pixels was used for normal imaging applications as described in chapter 3 with the photodiode free of wire-grid micro-polarizer while the other region was made polarization sensitive by using the wire-grid micro-polarizer on top of the photodiode.

The polarization analyzing region was further split into two regions of 64x64 arrays, referred to as sense region 1 and 2 respectively. A 2x2 pixel array in the sense region 1 consists of two pixels (A and B) measuring the intensity, while the other two measure the 0° (D), and 90° (C) polarized intensity respectively. A 2x2 pixel array in the sense region 2 has one pixel recording the intensity of the light (A) while the other three record the 0° (B), 45° (C) and 90° (D) polarized intensity.

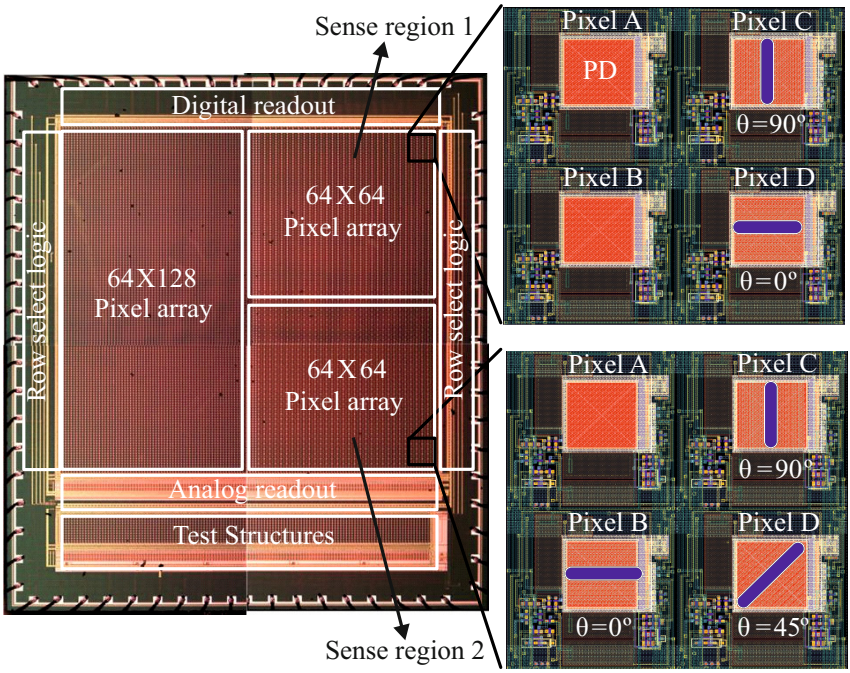
Figure 4.32(b) shows the transmission image showing the attenuation of the light in the regions where the photodiode has integrated metallic wire-grid polarizer. The sense region 2 has three pixels with wire-grid polarizer thus the attenuation of light is greater compared to sense region 1.

4.6 Polarization Measurements

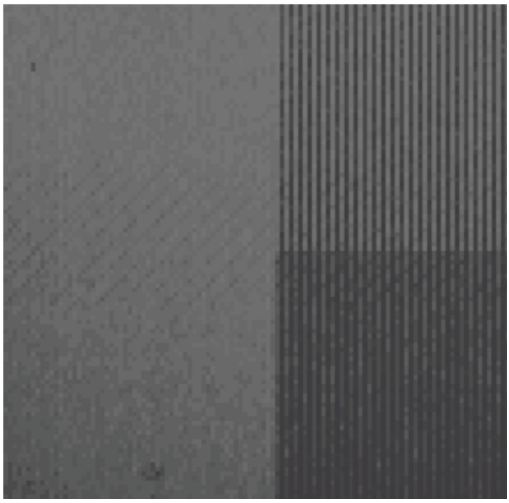
A partially linearly polarized light has three descriptors: intensity, degree of polarization and *e*-vector angle. These descriptors are measured using the variations in the pixel intensity with the changes in the transmission axis of the linear polarizers. The transmission efficiency is a measure of the partially linearly polarized light intensity that has been transmitted through the linear polarizer. The transmission efficiency, extinction ratio and degree of polarization of the designed wire-grid micro-polarizer are measured and are found to be fully functional as a linear polarizer.

4.6.1 Measurement Setup

The measurement setup for measuring the transmittance and the extinction ratio is shown in figure 4.33. In comparison to the measurement setup shown in the figure 3.40, here an additional linear external polarizer is used. The sensor was



(a)



(b)

Fig. 4.32 (a) Sensor regions with different polarizing angles, (b) transmission photograph of the wire-grid polarizer

illuminated with a polarized light obtained by passing the light from a DC light source through a linear polarizer. The DC light source used is an Philips halogen 12V, 20W bulb, set to radiate $400\mu W/cm^2$. The optical power was measured using the optical power meter TQ8210. The transmission axis of the linear polarizer was then varied from 0° to 180° in steps of 15° to vary the polarization angle of the light reaching the image sensor.

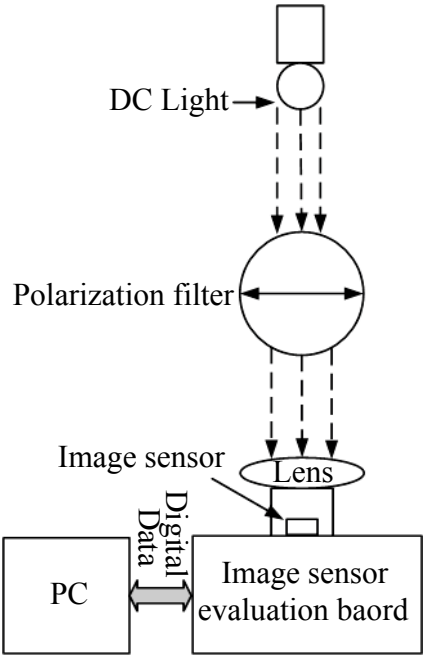


Fig. 4.33 Wire-grid measurement setup

The corresponding analog output of the pixels sensitive to 0° and 90° in the polarization sense region 1 and 0°, 45° and 90° in the polarization sense region 2 (see section 4.6) are recorded. The analog signal from the image sensor is digitized using an external *ADC* and then analyzed using a PC. The obtained output is normalized with respect to the intensity obtained at the intensity sensitive pixel, to compensate for any variation in the light intensity and pixel sensitivity over the entire imaging array. The normalized output is the transmittance of the wire-grid polarizer.

An average value of the intensities of the pixels sensitive to 0° and 90° in the sense region 1 and the pixels sensitive to 0°, 45° and 90° in the sense region 2 is computed for 30 frames using the following equation:

$$p_{avg(\text{with linear polarizer})}(x,y) = \frac{1}{N}\sum_{n=1}^{N} p_n(x,y) \tag{4.35}$$

where $p_n(x, y)$ is the measured pixel intensity in frame n, x and y correspond to the row and column address of the pixel in the sensor array and N is the number of frames selected. At the beginning of the experiment the mean of an array of 20x20 pixels without the linear polarizer is calculated as in equation (4.36). The result is used as a normalization factor.

$$p_{avg(\text{without linear polarizer})}(x, y) = \frac{1}{XYN} \sum_{n=1}^{N} \sum_{X=0}^{X-1} \sum_{Y=0}^{Y-1} p_n(x, y) \tag{4.36}$$

where X and Y are the pixel array dimensions and N is the total number of frames used to compute the mean. The normalized intensity is then obtained by dividing the mean pixel intensity with the linear polarizer (equation (4.35)) by the mean pixel intensity without the linear polarizer (equation (4.36)), as shown in equation (4.37)).

$$\text{Normalized } p_{avg}(x, y) = \frac{\frac{1}{N} \sum_{n=1}^{N} p_n(x, y)}{\frac{1}{XYN} \sum_{n=1}^{N} \sum_{X=0}^{X-1} \sum_{Y=0}^{Y-1} p_n(x, y)} \tag{4.37}$$

4.6.2 *Analog Polarization Measurements*

The polarization transmission path of the incident light ray in the figure 4.33 can be simplified as shown in the figure 4.34. The polarizer is the external polarization filter and the analyzer is the metallic wire-grid polarization filter patterned on top of the photodiodes of the designed polarization image sensor. When an unpolarized light wave with intensity I_0 is incident on the polarizer, the intensity of the transmitted light wave is I_1 .When this light wave is further incident on the analyzer the intensity I_2 is obtained after transmission, which is then detected by the detector.

In the polarization sense region 1 of the image sensor (figure 4.32a), to measure the 90° polarization behavior, the transmission axis of the polarizer was varied from

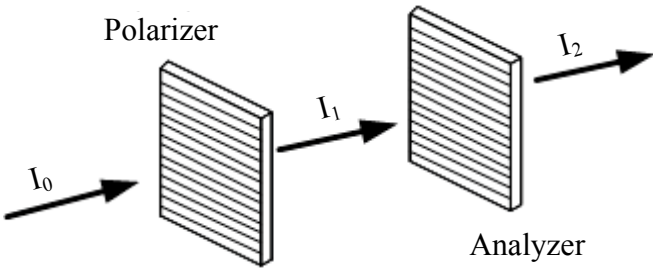


Fig. 4.34 Polarizer and analyzer arrangement

0° to 180° and the corresponding analog output of the pixel with a 90° transmission axis analyzer was measured. For the 0° polarization behavior, the transmission axis of the polarizer was varied from 0° to 180° and the corresponding analog output of the pixel with a 0° transmission axis analyzer was measured. The normalized transmittance as a function of the transmission axis of the linear polarizer (incident polarization angle) for the sense region 1 is shown in figure 4.35. The transmitted radiance of a linearly polarized light through the linear polarizer at a pixel array in an image is seen to vary sinusoidally as function of the polarizer transmission axis.

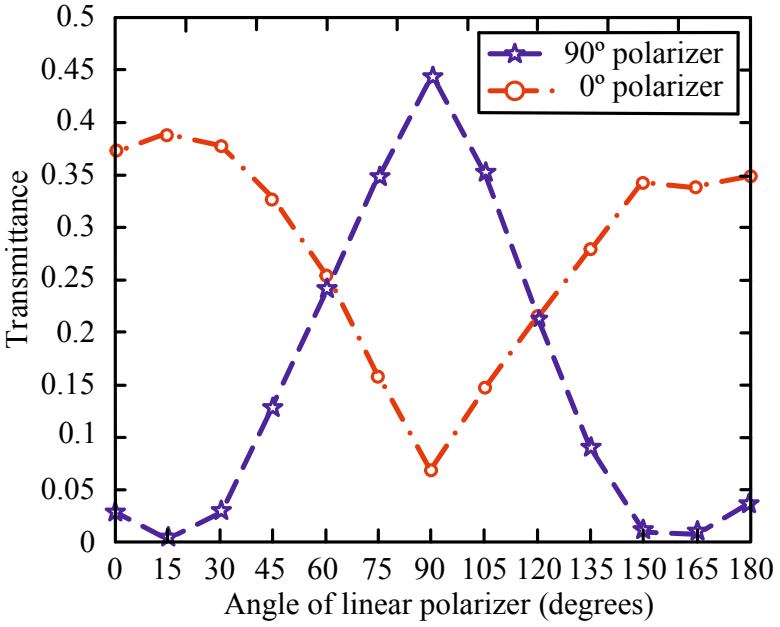


Fig. 4.35 Polarization sense region 1: 90° and 0° polarization transmittance

The maximum I_{max} of the sinusoid occurs when the polarizer orientation is parallel to the orientation of the linear polarized component and the minimum I_{min} occurs when the polarizer orientation is perpendicular to the linear polarized component. In the sense region 1, for 0° polarization the maximum transmittance achieved is 38.9% and the minimum transmittance observed is 7%. For 90° polarization the maximum transmittance obtained is 44.4% and the minimum transmittance is 0.1%.

The experiment was then repeated in the polarization sense region 2 of the image sensor (figure 4.32a). The normalized transmittance as a function of the transmission axis of the linear polarizer (incident polarization angle) for the sense region 2 is shown in figure 4.36. For the polarization sense region 2, for 0° polarization the

maximum transmittance obtained was 38.4% while the minimum transmittance was 5.45%. For 90° polarization the maximum transmittance observed was 42.4% while the minimum transmittance observed is 0.6%.

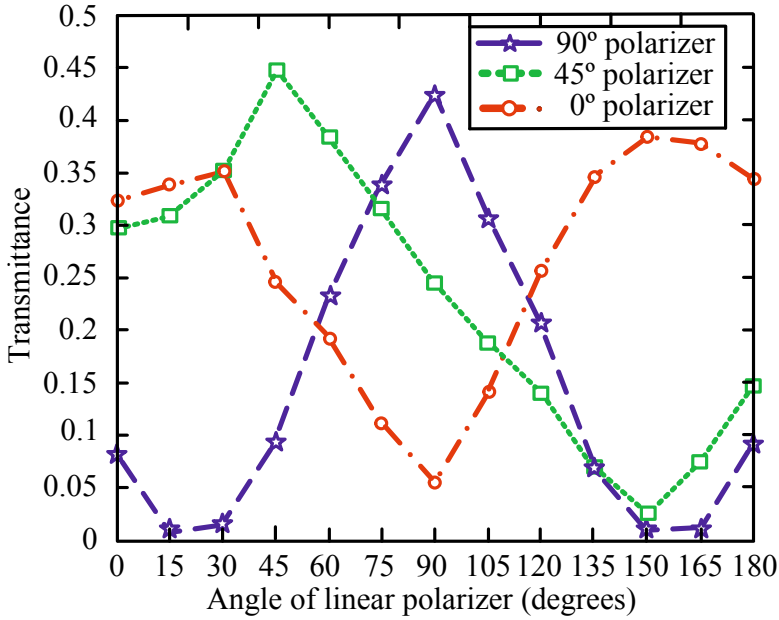


Fig. 4.36 Polarization sense region 2: 90°, 45° and 0° polarization transmittances

The extinction ratio can be calculated using equation (4.33). The calculated extinction ratio for the linear polarizer in the sense region 1 is 6.3 and in the sense region 2 is 7.7 [36] at λ=550nm. The obtained extinction ratios are higher than reported in literature for a similar polarization detection CMOS image sensor, where an extinction ratio of 2.3 is presented in [21], [22] with a 0.95μm wire-grid pitch designed in a 0.35μm 2-poly 4-metal standard CMOS at λ=589nm and 6 at λ=650nm in [37] using a photo-patterned micropolarizer array.

The absorption of the *EM* waves to completely *s*-polarize or *p*-polarize the transmitted wave through the wire-grid is dependent on the pitch of the wire-grid and so the extinction ratio also varies with the pitch. In our case, a higher extinction ratio is obtained than previously reported in literature because of the lower wire-grid pitch used in the image sensor. The mean maximum (T_{max}) and the minimum transmittance (T_{min}) for 0° and 90° polarization sensitive pixels in the polarization region 1 and 2 are shown in the table 4.3.

The sense region 2 has an additional linear polarizer oriented at 45°. The maximum and the minimum transmittance efficiency for the 45°sensitive pixel are 44.6% and 2%. The extinction ratio is thus calculated to be 22.

Table 4.3 Transmittance (%) and extinction ratio (ER)

Region	$T_{max}(0°)$	$T_{min}(0°)$	$T_{max}(90°)$	$T_{min}(90°)$	*ER*
1	0.389	0.07	0.444	0.01	6.3
2	0.384	0.0545	0.424	0.06	7.7

A performance comparison of the designed polarization sensor with the available wire-grid polarizers in literature is shown in table 4.4

Table 4.4 Performance comparison of the available polarization sensors

	[34]	[35]	[22]	This work
Imager type	CCD	CMOS	CMOS	CMOS
Polarizer type	Aluminum nanowires	Micro-patterning liquid crystal metal wire-polarizer	Standard CMOS metal layers	Standard CMOS metal layers
Spatial orientations	0°, 45°, 90° & 135 °	0°, 90° & right-handed circular polarized	0° & 90°	0°, 45° & 90°
Pixel pitch	7.4μm	5μm	20μm	25μm
Grating thickness	0.14μm	5μm	-	0.35μm
Grating width	0.75μm	-	0.6μm	0.24μm
Grating pitch	0.14μm	-	1.2μm	0.48μm
Extinction ratio	$\sim$60 (λ=530nm)	$\sim$1100 (λ=500nm)	2.3 (λ=633nm)	6.6 & 7.7 (λ=550nm)

The degree of polarization variation for the sense region 1 and 2 is shown in the figure 4.37. For a perfect polarizer I_s in equation (4.34) is equal to 1, which means full transmission of polarized light whose electric field is perpendicular to the wire-grid and I_p is equal to 0, meaning complete blockage of the polarized light whose electric field is parallel to the wire-grid.

The degree of polarization is shown to vary between 1 and -0.8. Ideally it is expected to vary between 1 and -1. The maximum value of the *DOP* is obtained at 0°and 180°while the minimum is obtained at 90°angle of the linear polarizer.

4.7 Wavelength Selection Using Metal Grid

The extinction ratio of a wire-grid polarizer depends on the incident wavelength and oblique angle incidence of the incident electromagnetic wave. The transmission decreases rapidly for the *s*-polarization with increasing wavelength or decreasing

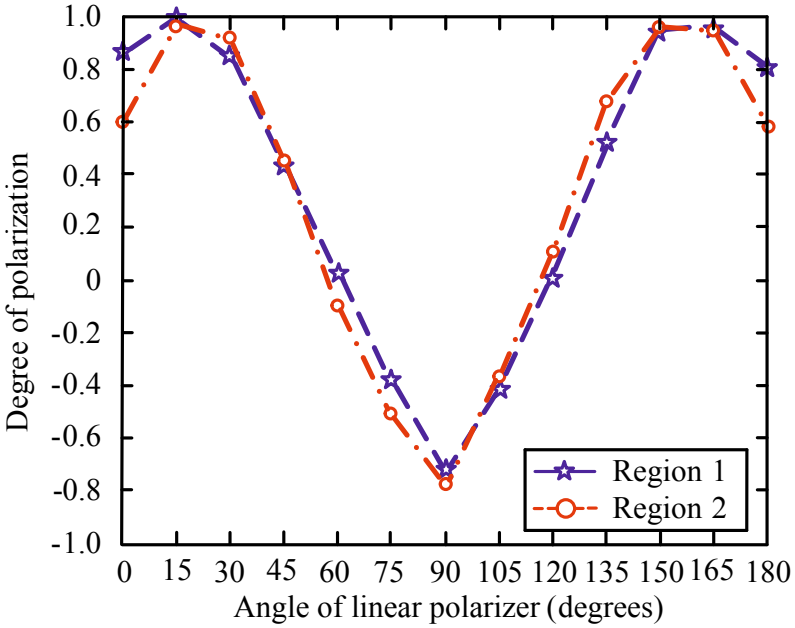


Fig. 4.37 Degree of polarization of the wire-grid

width of the opening between the metal wire-grid [38]. In contrast, the transmission of the *p*-polarized wave remains high for long wavelengths. The *ER* is found to be wavelength dependent: a low *ER* value is obtained for shorter wavelengths [38]. Transmission for *p*-polarization is at least two orders of magnitude higher than *s*-polarization in the 600-900nm range. To control the incident wavelength, regular metal grids have been used for a long time as filters, transmitting some wavelength and blocking others. The selective transmission of the wavelength can further be used as a color filter.

A novel technique for color detection using metal gratings on top of the photodetectors in standard CMOS technology was first introduced in 2001 [39]. The metal gratings are two-dimensional structures with different periodicity. Metal gratings have not been used for color filtering until recent years because the grid spacing must be very small for visible wavelengths; this has been difficult until 0.18μm CMOS technology became available [40]. The metal layers are deposited on top of the photodetectors, and since commercial IC fabrication processes have between two and eight layers of metal available, integrating metal grid light filters with the detectors eliminates the need for additional color filter array manufacturing steps. The short distance between the light filter and the photodetector has the advantage of reducing color crosstalk and pixel vignetting [39]. The light efficiency decreases due to the loss of light in the path from a light source to the pixel in the presence of wire-grid polarizer's. However for applications like for example *LCDs*

which contains two layers of polarizer's and a color sheet the light efficiency is very low, near 10%. The patterned metal gratings on the pixel will help to detect both the polarization and the color allowing for increased light efficiency compared to conventional *LCDs* type applications and also allow for multispectral imaging.

A simulation was performed to check the feasibility of selective transmittance of a desired wavelength (color filter) using one-dimension or two-dimensional metal patterns on top of the photodetector in a CMOS or CCD image sensor pixel. A dual layer aluminum grid with a width of 0.2μm and a spacing of 0.1μm is used as a basic structure to design a color filter as shown in figure 4.38.

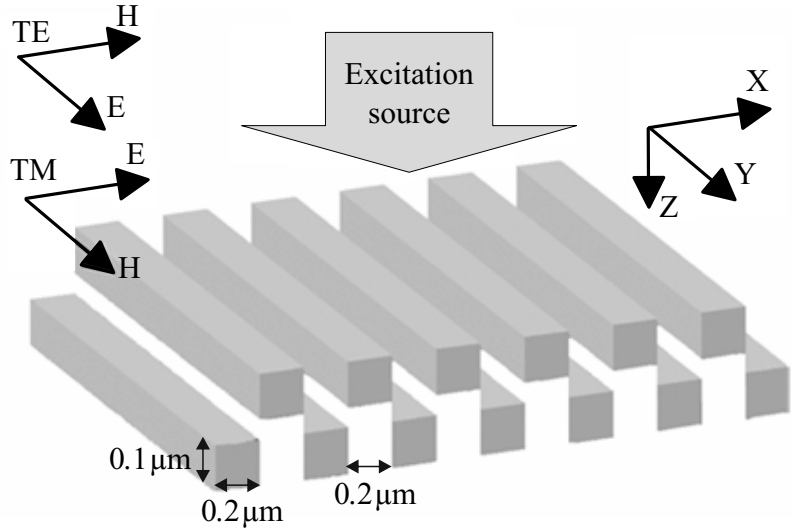


Fig. 4.38 2-layer wire-grid polarizer

A two-dimensional forward difference time domain *(FDTD)* method of solving the Maxwell's equations for the incident electromagnetic wave field is used to simplify the simulations. *FDTD* is a time-domain approach to numerically solve the Maxwell's equations by continuously solving the electric and magnetic fields of the incident electromagnetic field over time to simulate the propagation of the waves.

A broadband Gaussian-modulated pulsed light source expressed by equation (4.38) is used to illuminate the wire-grid filter along the positive z direction.

$$T(t) = \exp\left[-\frac{1}{2}\left(\frac{t - t_{off}}{t_w}\right)\right]\sin(\omega t) \tag{4.38}$$

where t_{off} is the offset time, t_w the half width of the pulse, and ω the central frequency of the source. Using this broadband excitation source, information of all the frequencies could be obtained in a single calculation. Due to the invariance of

the structure along the y-axis, the transverse-magnetic *(TM)* mode, consisting of the field components E_x, H_y and E_z and the transverse-electric *(TE)* mode, consisting of E_y, H_y and E_z are completely decoupled.

Figure 4.39 shows the transmittance and reflection of an excitation signal polarized in 0° and 90°, applied in the perpendicular *(z)* direction to the metal grid structure with 0° phase shift between the two metal layers (with the edge of the top and bottom metal layer coinciding). Figure 4.39(a) shows that for a 0° polarized incident signal, the wire-grid structure transmits all the three primary colors. However for 90° polarized incident signal, figure 4.39(b), the wire-grid filter shows a reduced transmission for wavelengths of 530nm and above, thus effectively working as a blue filter with transmission maximum between 400nm and 500nm.

The transmittance and reflection of the 0° and 90° polarized excitation signals through the metallic wire-grid when the top metal layer in figure 4.38 is shifted by 0.1μm in the x direction to have a 90° phase shift between the two metal layers (an edge distance of 0.1μm between the top and the bottom metal layers) are shown in the figure 4.40. In this arrangement, for the 0° polarized incident signal the maximum transmission peak occurs between 590nm to 620nm thus transmitting the red color of the visible spectrum while attenuating the green and the blue light. For 90° polarized incident signal the wavelengths above 530nm are completely blocked thus effectively blocking the red and the green color of the visible spectrum while transmitting the blue color.

The transmittance and reflection characteristics obtained for 0° and 90° polarized excitation signal through the metallic wire-grid when the top metal layer in figure 4.38 is further shifted by 0.1μm to obtain an 180° phase shift between the two metal layers (edge distance of 0.2μm between the top and the bottom metal layers) are shown in the figure 4.41. In this configuration both for the 0° and 90° polarized incident signal only the blue region of the visible spectrum is transmitted while the red and the green are effectively attenuated or blocked.

The results demonstrate transmittance and responsivities potentially suitable for color filtering. By controlling the wavelength responsivity in the pixel design process, the polarization information for the responsive wavelength can also be fixed. There is a loss of transmission efficiency due to the increased attenuation and thus this will not be suited for high resolution imaging but it would help in polarization imaging.

Besides depending on the wavelength of the incident light, the *ER* is also a function of the angle of incidence of the light ray. At a specular interface between a dielectric and absorbing material such as a metal or semiconductor, the angle of reflection of an incident light is equal but opposite to the angle of incidence, while the transmitted ray is refracted in accordance with Snell's law. The reflectance and transmittance are a function of the angle of incidence and the real and complex refractive indices of the materials. They are also functions of the polarization of the incident ray in accordance with the Fresnel's formulae. At the Brewster's angle, the reflected light is completely polarized perpendicular to the plane of incidence. Brewster's angle depends on the indices of refraction of the incident medium (n_1)

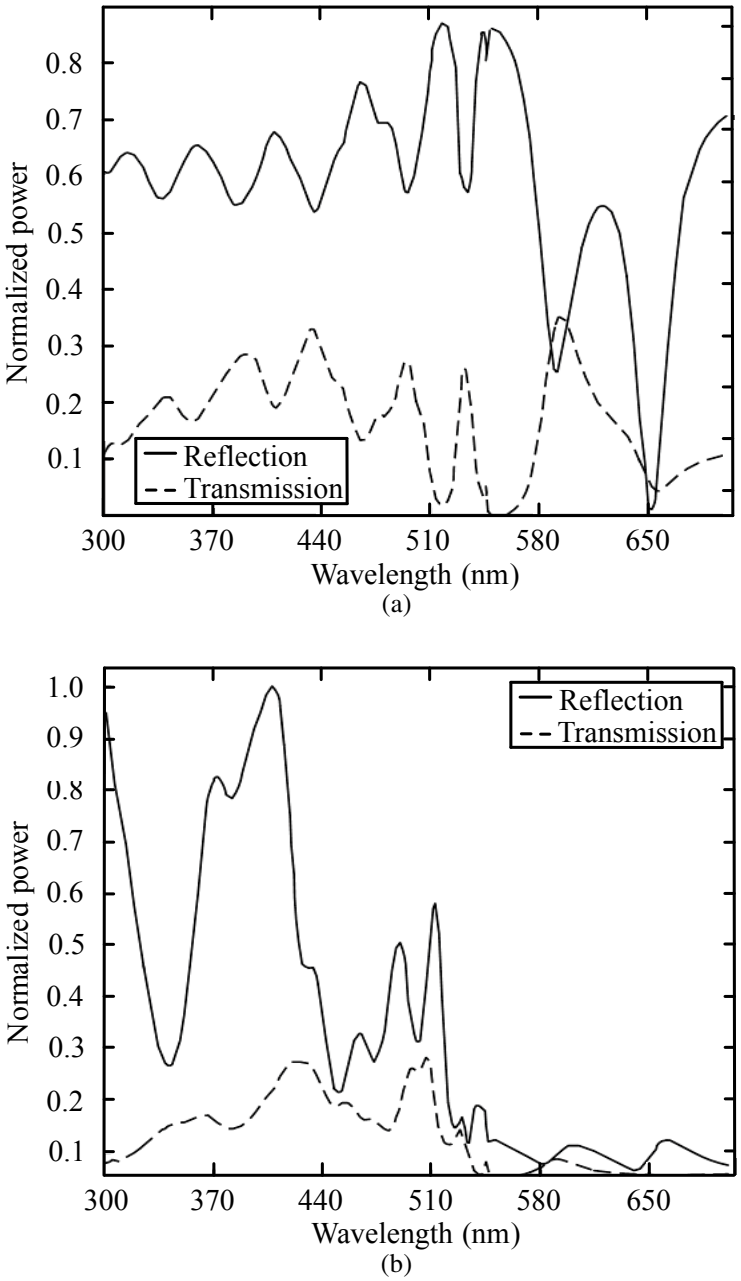


Fig. 4.39 0°phase shift between the metal layers, reflection and transmission of (a) 0° polarized signal b) 90° polarized signal

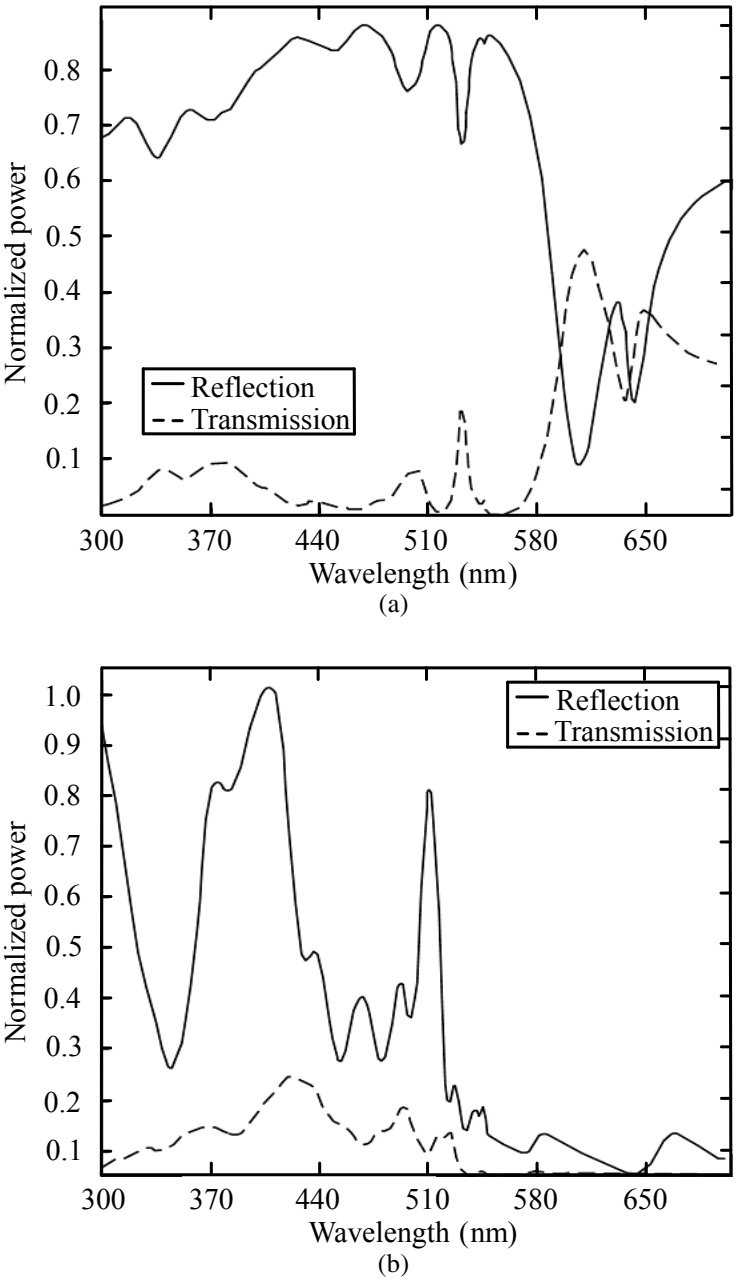


Fig. 4.40 90° phase shift between the metal layers, reflection and transmission of (a) 0° polarized signal b) 90° polarized signal

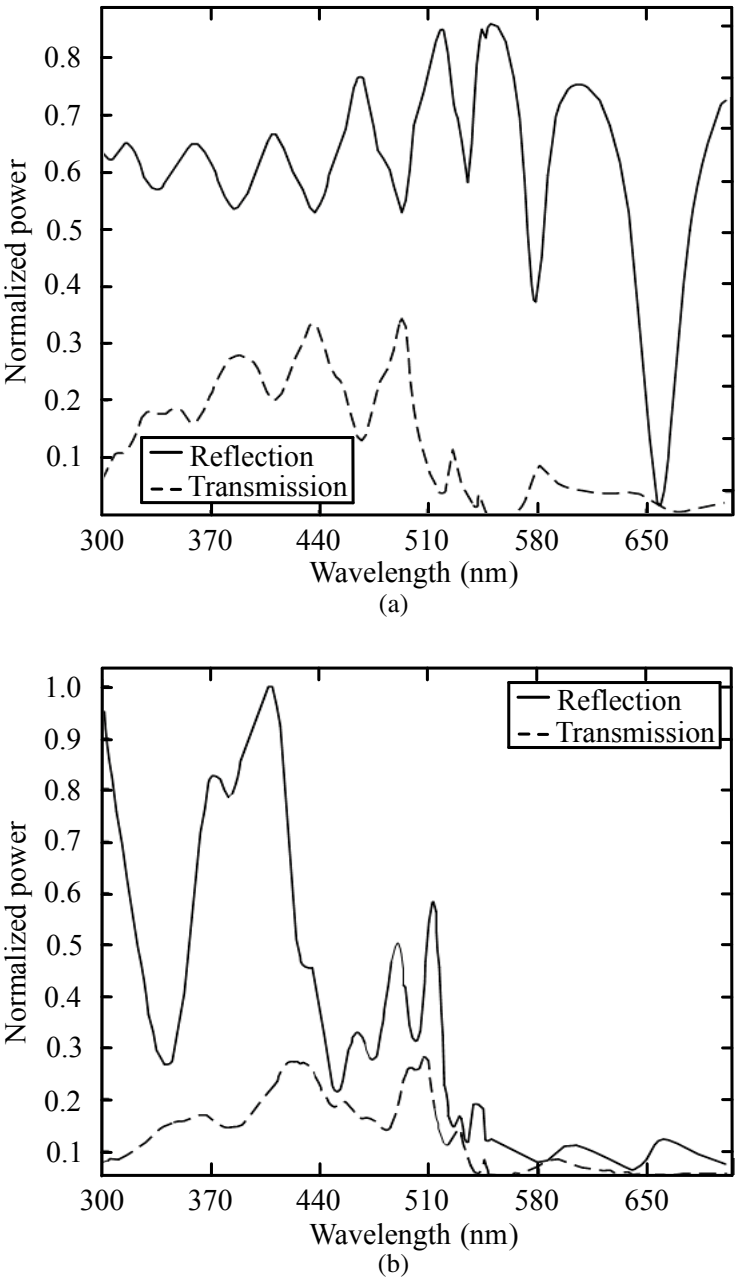


Fig. 4.41 180° phase shift between the metal layers, reflection and transmission of (a) 0° polarized signal b) 90° polarized signal

and the reflecting medium (n_2). A detailed review of the Fresnel's equations and Brewster's angle is provided in chapter 5.

The polarization effects increase with increasing angle of incidence of the incident light [41]. At angles greater than 0°, the component of lightwaves vibrating parallel to the plane of incidence and reflection (*p*-plane) will be filtered differently than the component vibrating perpendicular to the plane of incidence (*s*-plane). The reflectance of the *p*-polarization decreases with increasing angle of incidence to a minimum, at the quasi polarizing angle, and then increases towards one as the polarizing angle approaches 90°. The reflectance of the *s*-polarization monotonically increases towards one as the polarizing angle approaches 90°. The extinction ratio is shown to rapidly decrease for higher incident angles, for very high incident angles the polarization information is completely lost due to the pixel crosstalk [34].

Furthermore both polarization components also undergo a phase shift on reflection upon the angle of incidence. The *s*-polarization phase shift increases monotonically towards π radians and that of the *p*-polarization decreases monotonically towards zero as the polarizing angle approaches 90°. The change in the polarization behavior with changes in the angle of incidence affects the extinction ratios of the linear polarizers.

A wide field of view camera with a high angular resolution would be able to capture the polarization information over a wide field of view and this would naturally increase the sensitivity of the polarization or *ER* measurements.

4.8 Summary

- The three basic characteristics of light are intensity, wavelength or colour and polarization. Polarization is the plane of vibration of the electric field vector of light.
- Polarization is a phenomenon peculiar to transverse waves. It is the distribution of the electric field in the plane normal to the propagation direction.
- The reflected light in most environments is partially polarized, it can be represented by the sum of unpolarized states and completely linear polarized states.
- An unpolarized light can be polarized either by absorption, reflection, refraction or scattering.
- The state of the polarization can be completely defined by four parameters: two amplitudes, the magnitude and the sign of the phase shift.
- Stokes parameters are often used to define the state of the polarization. They describe the polarization state of light in terms of optical intensity. The Stokes parameters are obtained by measuring the transmitted intensity through a set of linear polarization filters.
- The degree of polarization *DOP* quantifies the fraction of the total optical signal that is actually polarized. The *DOP* of a fully polarized light is one, whereas the *DOP* of an unpolarized light is zero.

- Polarization filters are needed to visualize the polarization state of light. Polarization imaging cameras has been developed for visible, mid infra-red and long infra-red wavelength regions.
- The state-of-the-art polarization image sensors can be either operated in time domain multiplexed mode or in spatial domain multiplexed mode. In time domain multiplexed mode, the polarization filter is externally coupled to the photodetector array. In spatial domain multiplexed mode, the filters are fabricated directly on top of the photodetector.
- The micro-polarizer can either be directly fabricated on the sensor array or on a substrate and then flip-chip bonded with the sensor array. This reduces the spatial resolution of the obtained image, however the temporal resolution is increased compared to external filter.
- A wire-grid is a planar structure composed of a series of parallel wires or strips made of a conducting material. They act as diffraction gratings, however when the pitch is less than half of wavelength of the incident light, it act as a polarizer.
- The electric field parallel to the wire-grid (*TE* polarization) is reflected back where as the electric field perpendicular to the wire-grid (*TM* polarization) passes through the wire-grid.
- The key parameters that determine the performance of a wire-grid polarizer are its period (pitch), linewidth, and depth of the grid elements.
- Linear polarizers are characterized by transmission efficiency and extinction or contrast ratio.
- Aluminum is the preferred choice for fabricating wire-grid polarizer because it allows maximum p-polarization transmission and reflects most of the s-polarization.
- The extinction ratio of a wire-grid polarizer can be enhanced by considering two wire-grids with their transmission axis is series.
- With the scaling of the CMOS technology, the metal lines used for interconnections are scaling too and can be used as wire-grid polarizer. A wire-grid micro-polarizer was implemented using the standard metal layers of CMOS technology.
- The implemented micro-polarizer with CMOS metal layers showed polarization sensitivity, and extinction ratios of 6.3 and 7.7 were achieved. The extinction ratios of available external linear polarizers range from 10 to over 10^7.
- The extinction ratio is a function of the wire-grid pitch and with the scaling of the technology the pitch can be lowered and thus would help to achieve higher extinction ratios.
- The metal grid on top of the photodiode also affects the wavelength responsivity. Two-dimensional metal patterns show high wavelength selectivity but low transmittance.
- The use of metal grid severely attenuates the light and thus not suitable for resolution imaging, however they are ideal for machine vision applications.

References

[1] Bernard, G., Wehner, R.: Functional similarities between polarization vision and color vision. Vision Research 17, 1019–1028 (1977)

[2] Wehner, R.: Polarization vision-a uniform sensory capacity? Journal of Experimental Biology 204, 2589–2596 (2001)

[3] Reiley, D.: Polarization in optical design. PhD thesis, University of Alabama (1993)

[4] Horváth, G., Varjú, D.: Polarized light in animal vision. Springer, Heidelberg (2003)

[5] http://www.microscopyu.com/articles/polarized/index.html

[6] Stokes, G.: On the comppostion and resolution of streams of polarized light from different sources. Transactions of Cambridge Philosophical Society 9, 399–416 (1852); reprinted in Mathematics and Physics Papers, vol. 3, p. 233. Cambridge University Press (1901)

[7] Poincaré, H.: Theorie Mathematique de la Lumiere, ch. 12, vol. 2. Gauthiers-Villars (1892)

[8] Jerrard, H.: Transmission of light through birefringent and optically active media: the poincaré sphere. Journal of Optical Society of America 44(8), 634–640 (1954)

[9] Tyo, J., Goldstein, D., Chenault, D., Shaw, J.: A review of passive imaging polarimetery for remote sensing applications. Applied Optics 45, 5453–5469 (2006)

[10] November, L.J., Wilkins, L.M.: Liquid crystal polarimeter: a solid state imager for solar vector magnetic fields. Optical Engineering 34(6), 1659–1668 (1995)

[11] de Lang, H., Bouwhuis, G.: Colour separation in colour-television cameras. Philips Technical Review 24(9), 263–298 (1962)

[12] Murata, N., Hirano, C., Ohoba, M., Nagahara, S.: Development of a 3-mos color camera. SMPTE Journal, 1270–1273 (1983)

[13] Andreou, A., Kalayjian, Z.: Polarization imaging: principles and integrated polarimeters. IEEE Sensors Journal 2(6), 566–576 (2002)

[14] Zhao, X., Boussaid, F., Bermak, A., Chigrinov, V.: Thin photo-patterned micropolarizer array for cmos image sensors. IEEE Photonics Technology Letters 21(12), 805–807 (2009)

[15] Gruev, V., der Spiegel, J.V., Engheta, N.: Image sensor with focal plane polarization sensitivity. In: Proceedings of International Symposium on Circuits and Systems, pp. 1028–1031 (2008)

[16] Faris, S.: Methods for manufacturing micropolarizers, u.s. patent 5, 327, 285 (1994)

[17] Guo, J., Brady, D.: Fabrication of thin-film micropolarizer arrays for visible imaging polarimetry. Applied Optics 39(10), 1486–1492 (2000)

[18] Gruev, V., Ortu, A., Lazarus, N., der Spiegel, J.V., Engheta, N.: Fabrication of a dual-tier thin film micropolarization array. Optical Express 15(8), 4994–5007 (2007)

[19] Hertz, H.: Ueber die einwirkung einer geradlinigen electrischen schwingung auf eine benachbarte strombahn. Annalen der Physics 270(5), 155–170 (1888)

[20] Nordin, G., Meier, J., Deguzman, P., Jones, M.: Micropolarizer array for infrared imaging polarimetry. Journal of Optical Society of America, A 16, 1168–1174 (1999)

[21] Tokuda, T., Yamada, H., Shimohata, H., Sasagawa, K., Ohta, J.: Polarization-analyzing cmos image sensor with embedded wire-grid polarizer. In: Proceedings of International Image Sensor Workshop (2009)

[22] Tokuda, T., Yamada, H., Sasagawa, K., Ohta, J.: Polarization-analyzing cmos image sensor with monolithically embedded polarizer for microchemistry systems. IEEE Transactions on Biomedical Circuits and Systems 3(5), 259–266 (2009)

[23] Wang, J., Walters, F., Liu, X., Sciortino, P., Deng, X.: High-performance, large area, deep ultraviolet to infrared polarizers based on 40nm line/78nm space nanowire-grids. Applied Physics Letters 90, 061104–1–3 (2007)
[24] Ahn, S.W., Lee, K.D., Kim, J.S., Kim, S.H., Park, J.D., Lee, S.H., Yoon, P.W.: Fabrication of a 50 nm half-pitch wire grid polarizer using nanoimprint lithography. Nanotechnology 16, 1874–1877 (2005)
[25] Lamb, H.: On the reflection and transmission of electric waves by a metallic grating. Proceedings of London Mathematical Society 29(1), 523–546 (1898)
[26] Weber, T., Fuchs, H.J., Schmidt, H., Kley, E.B., A.T.
[27] Malacara, D.: Physical optics and light measurements, vol. 26, p. 157. Academic Press (1989) ISBN:0124759718
[28] Carter, J.M., Savas, R.C., Sakes, T.A., Walsh, M.E., O'Reilly, T.: Interfernce lithography. MTL Annual Report, Submicron and Nanometer Structures, 186–188 (2003)
[29] Chen, C., Yu, T., Lee, Y.: Direct metal transfer lithography for fabricating wire grid polarizer on flexible plastic substrate. Journal of Microelecttromechanical Systems 20(4), 916–921 (2011)
[30] Collett, E.: Field Guide to Polarization, p. 25. SPIE Press, Bellingham (2005) ISBN: 9780819458681
[31] Zhou, Y., Klotzkin, D.J.: Design and parallel fabrication of wire-grid polarization arrays for polarization-resolved imaging at 1:5μm. Applied Optics 47(20), 3555–3560 (2008)
[32] Hu, C., Liu, D.: Investigation of aluminum wire-grid polarizers for visible wavelengths using rigorous coupled wave analysis. In: Proceedings of SPIE, vol. 7134 (2008)
[33] Xu, M., Urbach, H.P., de Boer, D.K.G., Cornelissen, H.J.: Wire-grid diffraction gratings used as polarizing beam splitter for visible light and applied in liquid crystal on silicon. Optics Express 13(7), 2303–2320 (2005)
[34] Gruev, V., Perkins, R., York, T.: Ccd polarization imaging sensor with aluminum nanowire optical filters. Optics Express 18, 19087–19094 (2010)
[35] Zhao, X., Bermak, A., Boussaid, F., Chigrinov, V.: Liquid-crytal micropolarimeter array for visible linear and circular polarization imaging. In: Proceedings of International Symposium on Circuits and Systems, pp. 637–640 (2009)
[36] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.J.P.: Polarization analyzing cmos image sensor. In: Proceedings of International Symposium on Circuits and Systems, pp. 621–624 (2010)
[37] Zhao, X., Bermak, A., Boussaid, F.: A cmos digital pixel sensor with photopatterened micropolarizer array for real time focal plane polarization imaging. In: Proceedings of IEEE Biomedical Circuits and Systems Conference, pp. 145–148 (2008)
[38] Guillaumee, M., Dunbar, L., Santschi, C., Grenet, E., Eckert, R., Martin, O., Stanley, R.: Polarization sensitive silicon photodiodes using nano-structured metallic grids. Applied Physics Letters 94, 193503-1–193503-3 (2009)
[39] Catrysse, P., Wandell, B.: Integrated color pixels in 0.18μm complementary metal oxide semiconductor technology. Journal of Optical Society of America A 20(12), 2293–2306 (2003)
[40] Catrysse, P., Wandell, B., Gamal, A.E.: An integrated color pixel in 0.18μm cmos technology. Technical Digest of IEEE International Electron Device Meeting, 24.4.4 (2001)
[41] Goldstein, D., Chenault, D., Pezzaniti, J.: Polarimetric characterization of spectralon. In: Polarization An Integrated Color Pixel in 0.18μm CMOS Technology: Proceedings of Measurement, Analysis, and Remote Sensing II SPIE, vol. 3754, pp. 126–136 (1999)

Chapter 5
Material Classification Using CMOS Polarization Sensor

Material classification is an important application in computer vision. The ability to detect the nature of the object surface from image data has a very high potential for applications ranging from low-level inspection to high-level object recognition. The inherent property of materials to partially polarize the reflected light can serve as a tool to classify them.

In this chapter, an introduction to material classification is presented in section 5.1. The theory behind polarization upon specular reflection and Fresnel coefficients is covered in section 5.2. The Fresnel reflectance model is also briefly described. The polarized nature of the transmitted light can be used to differentiate between metal and dielectric surfaces in real time due to the different nature in partially polarizing the specular and diffuse reflection components of the reflected light. This is elaborated in section 5.3 by measuring the transmitted irradiance after reflection from the material surface which allows computation of Fresnel reflection coefficients, the degree of partial polarization and the variations in the maximum and minimum transmitted intensities for varying specular angle of incidence. Differences in the physical parameters for various metal surfaces result in different surface reflection behavior, influencing the Fresnel reflection coefficients. Section 5.4, shows that it is possible to differentiate among various metals of varying conductivity by sensing the change in the Polarization Fresnel Ratio and the degree of polarization of the light reflected. Section 5.5 presents a short summary of the methods used to differentiate among material surfaces.

5.1 Introduction

Materials can be broadly classified into metals and dielectrics based on their conductivity. Metals are highly conductive, opaque and tend to be very reflective while dielectrics are less conductive and have very low reflectivity.

Earlier attempts to distinguish between metals and dielectrics used the dichromatic reflection properties of the material surface [1], [2], [3]. Materials were classified into optically homogenous or optically inhomogeneous categories.

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 157–184.
DOI: 10.1007/ 978-3-642-34901-0_5 © Springer-Verlag Berlin Heidelberg 2013

Healey [4] showed that homogenous materials reflect light only from the surface. Thus, when illuminated with a monochromatic light beam the reflected light color would be nearly constant. Inhomogeneous materials on the other hand reflect light from the surface and also scatter light from the body. Such a material, when illuminated with a monochromatic light beam would reflect two distinct colors as the color of the surface reflected light will be different from the color of the scattered light from the body.

Once the object has been determined to be either homogenous or inhomogeneous, it can be further classified as a metal or a dielectric depending on the reflected color. Metals have free electrons, and the behavior of these free electrons when hit by an incident light ray is a function of its wavelength. Metals such as steel and nickel have nearly the same response over the entire visible spectrum, while metals such as copper and gold tend to reflect longer wavelengths more than the shorter ones. Dielectrics on the other hand do not have free electrons thus the reflectivity from a dielectric surface is independent of the wavelength of the incident light ray. This method, however, is limited by the distortion of the color histogram obtained after reflection, which depends on the object geometry. There is also speculation whether a definitive relationship exists between the reflected color distribution and the intrinsic composition of the surface. Additionally, the color of the material also influences the reflected color.

Another method to discriminate between metals and dielectrics based on the Fresnel reflection theory was proposed by Wolff [5], [6]. An unpolarized electromagnetic wave can be polarized upon specular reflection as described in chapter 4. The polarization by reflection from a material surface depends on the angle at which the light strikes the reflecting surface as well as on the nature of said surface. Metallic surfaces reflect light with a variety of vibration directions, and such reflected light is usually unpolarized. However, non-metallic surfaces reflect light such that the vibrations of the reflecting light wave is parallel to the plane of the reflecting surface.

According to the Fresnel reflection theory, dielectric surfaces polarize the light upon specular reflection stronger than metal surfaces for all angles of incidence. Traditionally the polarization components of the transmitted irradiances are obtained by allowing the reflected light from the material surface to pass through an external linear polarizer onto a CCD or CMOS image sensor. The disadvantage of such a system is that the linear polarization filters have to be externally controlled, which complicates the automation and miniaturization of optical sensors for material classification. Further the digital processing required to process the obtained polarization information increases the complexity and power consumption of the system.

The model proposed by Wolff to attempt to classify materials based on the polarization information using the embedded wire grid metallic micro-polarizer presented in chapter 4 is used in this book. The metallic wire grid micro-polarizer can be used to measure the transmitted irradiances after specular reflection from the material surface in real time. The metal grid oriented at 0° and 90° in the sense regions 1 and 2 allows us to measure the parallel and perpendicular Fresnel

reflection coefficients and the maximum and the minimum transmitted irradiance after reflection from the material surface. The difference between the maximum and the minimum transmitted radiances relates to the degree of polarization at the surface [7]. Additionally, the degree of polarization based on the Stokes parameters discussed in chapter 4 is evaluated to find the relative difference between metals and dielectrics.

5.2 Polarization and Fresnel Coefficients

When a light ray strikes a surface, part of the incident light is reflected and part is transmitted or absorbed as shown in figure 5.1. φ is the angle of incidence of the incident ray, φ' is the angle of reflection of the reflected ray and η is the angle of transmission of the transmitted ray.

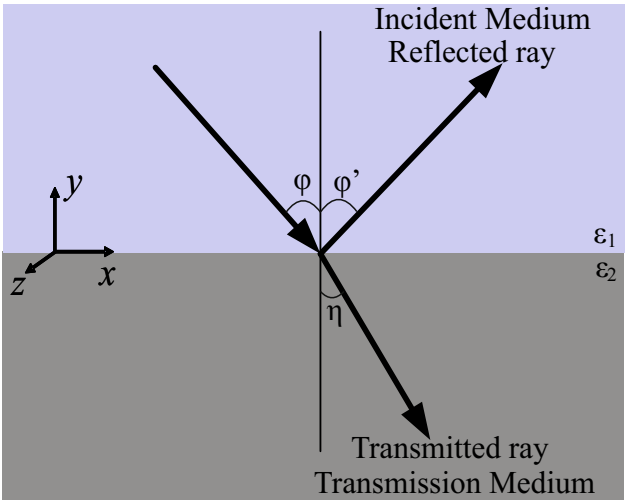


Fig. 5.1 Incident light ray behavior at an interface

The incident ray and the reflected ray are related by the law of reflection:

$$\varphi = \varphi' \tag{5.1}$$

While the incident ray and the transmitted rays are related by the law of refraction or Snell's law:

$$\epsilon_1 \sin\varphi = \epsilon_1 \sin\eta \tag{5.2}$$

where ϵ_1 is the refractive index of the incident medium and ϵ_2 is the refractive index of the transmission medium.

The reflection occurring at the surface of the planar surface has two components: diffuse reflection and specular reflection, as shown in the figure 5.2.

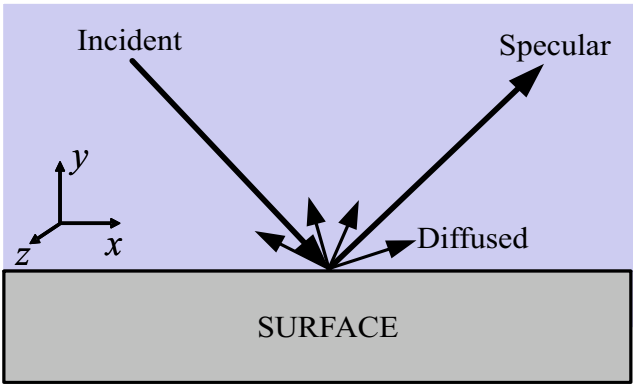


Fig. 5.2 Reflection types: specular and diffuse reflections

Diffuse reflection is caused by the reflected rays from internal scattering inside the surface of the medium. When light strikes a surface interface, a part of the light passes through the boundary. This light is re-emitted randomly after suffering internal scattering. This re-emitted light is known as diffuse reflection and is modeled by Lambert's equation:

$$\begin{aligned} I_d &= (C_d \times \vec{N}).\vec{S} \\ &= (C_d \times \cos\theta_i) \end{aligned} \tag{5.3}$$

where I_d corresponds to the brightness C_d is a proportionality constant, $\vec{N}$ is the surface orientation, $\vec{S}$ is the light source direction and θ_i (like φ in figure 5.1) is the angle between the light source direction and the surface orientation. The diffuse reflection component is independent of the angle of reflection but depends on the angle of incidence.

Specular reflection is a mirror like reflection from a surface in which light from a single incoming direction is reflected into a single outgoing direction. Pure specular reflection occurs when the planar interface portion of the surface is significantly larger than the wavelength of the incident light [5]. The incident and the reflected directions of the specularly reflected light determine the specular plane of incidence (or simply specular plane) as shown in figure 5.3.

5.2.1 Polarization Properties of a Reflected Light

Unpolarized light becomes partially polarized after specular reflection as shown in figure 5.4. In figure 5.4 the plane of incidence for a plane wave is *xz*, the plane of the two mediums is *xy* and the angle of incidence is φ. For the reflected light, the intensity of the polarization component perpendicular to the specular plane is larger than the intensity of the polarization component parallel to the specular plane.

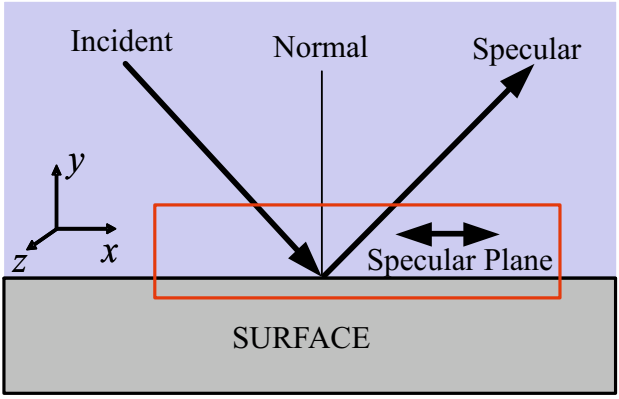


Fig. 5.3 Specular reflection plane is parallel to x-y plane

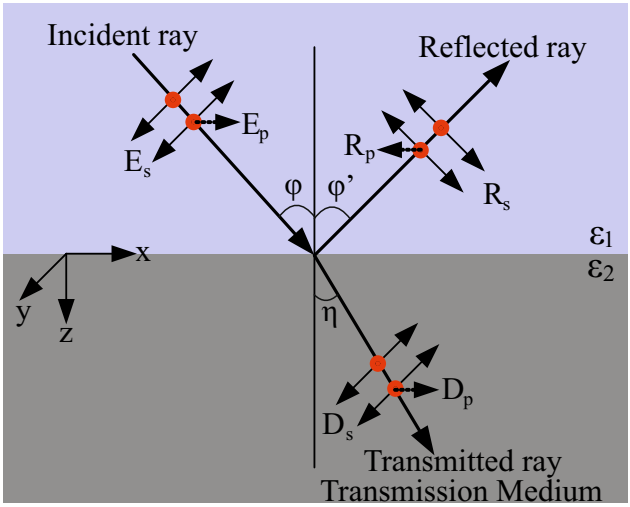


Fig. 5.4 Incident, reflected and transmitted Fresnel coefficients

A planar wave is represented by the equations:

$$
\begin{aligned}
X &= A_x \cos \frac{2\pi}{T}(t - \frac{mx + ny + pz}{V}) \\
Y &= A_y \cos \frac{2\pi}{T}(t - \frac{mx + ny + pz}{V}) \\
Z &= A_z \cos \frac{2\pi}{T}(t - \frac{mx + ny + pz}{V})
\end{aligned}
\tag{5.4}
$$

where A_x, A_y and A_z are the components of the amplitude of the resultant electric forces in the *x*, *y* and *z* directions, respectively; *V* is the velocity of propagation; *m*, *n* and *z* are the direction cosines of the normal to the wave front and fulfill the equation $m^2 + n^2 + p^2 = 1$; *T* is the time period of the wave and *t* is the time independent variable [8].

The phase in equation (5.4) is the same for all planes thus the wave fronts can be written as:

$$mx + ny + pz = constant \tag{5.5}$$

The amplitudes of the resultant electric forces, A_x, A_y and A_z are proportional to the directional cosines and thus satisfy equation (5.6).

$$(A_x \times m) + (A_y \times n) + (A_z \times p) = 0 \tag{5.6}$$

The direction cosines of the direction of propagation of the incident wave in figure 5.4 are then expressed as:

$$\begin{aligned} m &= \sin\varphi \\ n &= 0 \\ p &= \cos\varphi \end{aligned} \tag{5.7}$$

The incident electric field can be resolved into two components, one perpendicular to the plane of incidence with amplitude E_s and the other parallel to the plane of incidence or in the plane of incidence with amplitude E_p. The first component is parallel to the *y*-axis and can be expressed using equation (5.4) and (5.7) as:

$$Y_e = E_s \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi + z\cos\varphi}{V_1}) \tag{5.8}$$

V_1 is the velocity of the light in the incident medium:

$$V_1 = \frac{c}{\sqrt{\epsilon_1}} \tag{5.9}$$

where, *c* is the speed of light and ϵ_1 is the index of refraction of the incident medium.

The incident wave is transverse in nature with an electric force E_p perpendicular to the plane of incidence, thus the components A_x and A_z in the *x* and *z* directions must have the values:

$$\begin{aligned} A_x &= E_p\cos\varphi \\ A_z &= -E_p\sin\varphi \end{aligned} \tag{5.10}$$

The *x* and *z* components of the electrical force of the incident wave are then expressed as:

$$X_e = E_p \cos\varphi \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi + z\cos\varphi}{V_1})$$
$$Z_e = -E_p \sin\varphi \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi + z\cos\varphi}{V_1}) \tag{5.11}$$

When the incident transverse wave reaches the boundary it is divided into a reflected and a refracted wave component. The electric forces in the reflected wave are:

$$X_r = R_p \cos\varphi' \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi' + z\cos\varphi'}{V_1})$$
$$Y_r = R_s \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi' + z\cos\varphi'}{V_1}) \tag{5.12}$$
$$Z_r = -R_p \sin\varphi' \cos\frac{2\pi}{T}(t - \frac{x\sin\varphi' + z\cos\varphi'}{V_1})$$

And the electric forces in the refracted wave are:

$$X_2 = D_p \cos\eta \cos\frac{2\pi}{T}(t - \frac{x\sin\eta + z\cos\eta}{V_2})$$
$$Y_2 = D_s \cos\frac{2\pi}{T}(t - \frac{x\sin\eta + z\cos\eta}{V_2}) \tag{5.13}$$
$$Z_2 = -D_p \sin\eta \cos\frac{2\pi}{T}(t - \frac{x\sin\eta + z\cos\eta}{V_2})$$

where R_p, R_s, D_p and D_s are the amplitudes of reflected and refracted electrical forces, φ' is the angle of reflection, η is the angle of refraction and V_1 and V_2 are the velocity of the wave in the incident and the transmission medium respectively.

The boundary condition states that the component of the electric and magnetic forces parallel to the surface must vary continuously in passing through the transition layer. From equations (5.8), (5.11), (5.12) and (5.13) and using the law of reflection from equation (5.1) we obtain:

$$\begin{aligned}(E_p - R_p)\cos\varphi &= D_p \cos\eta \\ (E_s + R_s) &= D_s \\ (E_s - R_s)\sqrt{\epsilon_1}\cos\varphi &= D_s\sqrt{\epsilon_2}\cos\eta \\ (E_p + R_p)\sqrt{\epsilon_1} &= D_p\sqrt{\epsilon_2}\end{aligned} \tag{5.14}$$

From the above equations the amplitude of the reflected and refracted amplitudes can be calculated in terms of the incident amplitude as:

$$
\begin{aligned}
2E_s &= D_s(1 + \frac{\sqrt{\epsilon_2}\cos\eta}{\sqrt{\epsilon_1}\cos\varphi}) \\
E_s(\frac{\sqrt{\epsilon_1}\cos\varphi}{\sqrt{\epsilon_2}\cos\eta} - 1) &= R_s(\frac{\sqrt{\epsilon_1}\cos\varphi}{\sqrt{\epsilon_2}\cos\eta} + 1) \\
2E_p &= D_p(\frac{\cos\eta}{\cos\varphi} + \frac{\sqrt{\epsilon_2}}{\sqrt{\epsilon_1}}) \\
E_p(\frac{\cos\varphi}{\cos\eta} - \frac{\sqrt{\epsilon_1}}{\sqrt{\epsilon_2}}) &= R_p(\frac{\cos\varphi}{\cos\eta} + \frac{\sqrt{\epsilon_1}}{\sqrt{\epsilon_2}})
\end{aligned}
\tag{5.15}
$$

From the laws of refraction:

$$
\sqrt{\epsilon_1}\sin\eta = \sqrt{\epsilon_2} \tag{5.16}
$$

From equations (5.15) and (5.16) the Fresnel reflection equations can be derived:

$$
\begin{aligned}
R_s &= -E_s\frac{\sin(\varphi-\eta)}{sin(\varphi+\eta)} \\
R_p &= -E_p\frac{\tan(\varphi-\eta)}{tan(\varphi+\eta)} \\
D_s &= -E_s\frac{2\sin\eta\cos\varphi}{sin(\varphi+\eta)} \\
D_p &= -E_p\frac{2\sin\eta\cos\varphi}{sin(\varphi+\eta)cos(\varphi-\eta)}
\end{aligned}
\tag{5.17}
$$

where R_s and R_p represent the perpendicular and parallel Fresnel coefficients for reflection with D_s and D_p are the perpendicular and parallel Fresnel coefficients for transmission of the incident light ray.

From equations (5.17) it can be said that for any given incident ray there exists a state of polarization in the reflected and transmitted ray. The state of polarization of a light wave describes the electric field magnitude and phase relationships inside the reflected and the transmitted wave.

5.2.2 Fresnel Reflectance Model

The Fresnel reflectance model describes the intensity and spectral composition of the light reflected from the reflection surface and reaching the observer. The intensity of the reflected light depends on the intensity of the light source and also on the surface properties of the material. The spectral composition of the reflected light is determined by the wavelength selective reflection of the surface. In this section fresnel reflection coefficients are determined from the spectral composition of the reflected light which is further used to define polarization Fresnel ratio *(PFR)*.

When the reflected wave component is passed through a linear polarizer, the intensity of the image can be expressed as a function of the transmission axis of the

polarizer (θ) [5]. The spectral transmittances of the polarizer for linearly polarized light and unpolarized light are denoted by $T_p(\lambda)$ and $T_n(\lambda)$ respectively, where λ is the wavelength of the light.

The intensity image of the reflected light from the surface is expressed in terms of the pixel coordinates *(x,y)* and the transmission axis of the polarizer. The intensity image obtained after the reflection can be written as a sum of the diffuse reflection coefficient $I_d(x, y)$ and the specular reflection coefficient $I_s(x, y)$ [4.5],[4.7]. Using the Fresnel reflection equations (5.17), the intensity image observed through the polarizer transmission axis θ can be expressed as:

$$I(x, y : \theta) = T_n I_d(x, y) + T_p I_s(x, y) \times \frac{R_p(x, y)\sin^2(\theta - \theta_0) + R_s(x, y)\cos^2(\theta - \theta_0)}{R_p(x, y) + R_s(x, y)} \tag{5.18}$$

where θ_0 is the direction perpendicular to the specular plane. The Fresnel reflection coefficients R_p and R_s depend on the pixel coordinates. In terms of polarized and unpolarized components, equation (5.18) can be rewritten as:

$$I(x, y : \theta) = A(x, y)\frac{1 + \cos(2(\theta - \theta_0))}{2} + B(x, y) \tag{5.19}$$

where:

$$A(x, y) = \frac{R_p(x, y) - R_s(x, y)}{R_p(x, y) + R_s(x, y)} T_p I_s(x, y) \tag{5.20}$$

$$B(x, y) = \frac{R_p(x, y)}{R_p(x, y) + R_s(x, y)} T_p I_s(x, y) + T_n I_d(x, y) \tag{5.21}$$

The first term in equation (5.19) is the polarized component of the reflection mainly contributed by the specular reflection of the incident transverse wave. The second term represents the unpolarized component contributed by the diffuse reflections and partly by the specular reflection which is not polarized.

It can be observed from equation (5.19) that $I(x, y : \theta)$ oscillates as θ varies between the maximum $I_{max}(x, y)$ at $\theta = \theta_0, \theta_0 + \pi$ and the minimum $I_{min}(x, y)$ at $\theta = \theta_0 \pm \pi/2$. $I_{max}(x, y)$ and $I_{min}(x, y)$ are then expressed in terms of equations (5.20) and (5.21) as:

$$\begin{aligned} I_{max}(x, y) &= A(x, y) + B(x, y) \\ &= \frac{R_s(x, y)}{R_p(x, y) + R_s(x, y)} \times T_p I_s(x, y) + T_n I_d(x, y) \end{aligned} \tag{5.22}$$

$$I_{min}(x,y) = B(x,y) = \frac{R_{ps}(x,y)}{R_p(x,y)+R_s(x,y)} \times T_p I_s(x,y) + T_n I_d(x,y) \tag{5.23}$$

The Polarization Fresnel Ratio *(PFR)* is the ratio of the perpendicular Fresnel coefficient to the parallel Fresnel coefficient [5]. The Fresnel reflection and transmission coefficients in the Fresnel reflectance model are given by equation (5.17).

The *PFR* can be derived by rearranging the equations (5.22) and (5.23).

$$R_s(x,y) = \frac{[I_{max}(x,y) - T_n I_d(x,y)][R_p(x,y)+R_s(x,y)]}{T_p I_s(x,y)} \tag{5.24}$$

$$R_p(x,y) = \frac{[I_{min}(x,y) - T_n I_d(x,y)][R_p(x,y)+R_s(x,y)]}{T_p I_s(x,y)} \tag{5.25}$$

Dividing equations (5.24) and (5.25), we get

$$PFR = \frac{R_s(x,y)}{R_p(x,y)} = \frac{[I_{max}(x,y) - T_n I_d(x,y)]}{[I_{min}(x,y) - T_n I_d(x,y)]} \tag{5.26}$$

The specular component of reflection in metals is greater than the diffuse component of reflection $I_s(x,y) \gg I_d(x,y)$. Thus the diffuse component of reflection $I_d(x,y)$ in equation (5.26) can be neglected resulting in equation (5.27).

$$PFR = \frac{R_s(x,y)}{R_p(x,y)} = \frac{[I_{max}(x,y)]}{[I_{min}(x,y)]} \tag{5.27}$$

Equation (5.27) is the polarization Fresnel ratio which is related to the ratio of the maximum to minimum transmitted irradiance detected after specular reflection from the material surface.

In the case of dielectrics the Fresnel reflection coefficient R_p is very small for all specular angles of incidence, and is almost zero near the Brewster angle. The *PFR* for dielectrics thus can be arbitrary large as predicted by equation (5.27). For dielectrics with specular angle of incidence very close to the Brewster angle, equation (5.27) becomes

$$\frac{R_s(x,y)}{R_p(x,y)} \gg \frac{[I_{max}(x,y)]}{[I_{min}(x,y)]} \tag{5.28}$$

5.3 Material Classification Measurements

In the section 5.2, it was shown that the reflected and transmitted light polarization components are related to the incident polarization components. The reflected

polarization components are dependent on the angle of incidence and the nature of the material surface. In this section the variation in the polarization patterns for two materials, metal and plastic, are experimentally detected. The varying oscillations of the transmitted intensities, the degree of polarization and the *PFR* obtained from the polarization component of the reflected light for both the material surface are evaluated.

5.3.1 Measurement Setup

The measurement setup for the measurements of the maximum and minimum transmitted intensities after reflection from the material surface is shown in figure 5.5.

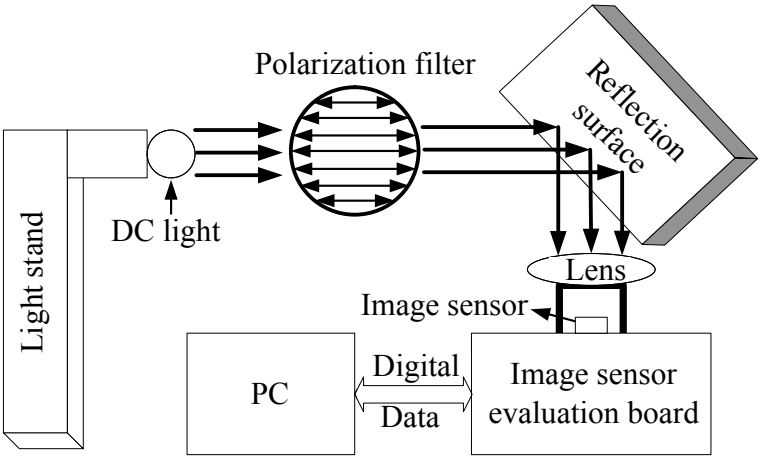


Fig. 5.5 Measurement setup for at λ=650nm

The light from the DC light source (similar to one used in 4.33) is completely unpolarized. The linear polarization filter allows only those amplitudes which are parallel to the transmission axis of the polarization filter to pass through. The polarized electromagnetic waves are reflected from the surface. At the boundary of the reflection surface both the diffuse component and the specular component of the reflection of the incident light are present. These reflection components are then focused on the image sensor by the lens. The analog signal from the image sensor is digitized using an external *ADC* and then analyzed using a PC as described in section 3.7.1.

5.3.2 Polarization Transmittance

The magnitude of variations of the reflected irradiance of the light after reflection is larger for dielectrics than for metals [5]. The theoretically expected image intensity

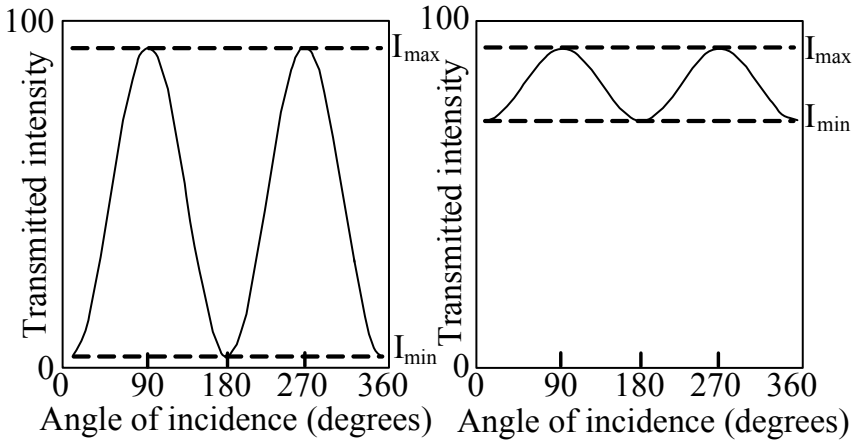


Fig. 5.6 Transmitted radiance of reflected light as a function of transmission axis of the linear polarizer (left) dielectric surface (right) metal surface

$I(x, y : \theta)$ of the transmitted radiance of light reflected from the dielectric and metal surfaces is shown in figure 5.6 [7].

The intensity of the reflected wave is a function of the transmission axis of the linear polarizer $I(x, y : \theta)$ and oscillates between the maximum transmitted intensity I_{max} and the minimum transmitted intensity I_{min} as shown in equation (5.19). For dielectrics the Fresnel coefficients satisfy $Rp \gg R_s$ while for metals $R_p \approx R_s$. The diffuse component of reflection dominates over the specular component $I_d(x, y) \gg I_s(x, y)$ for dielectrics and thus the oscillations given by equations (5.22) and (5.23) vary over a larger range, while in the case of metals, the oscillations are relatively smaller as the specular component of reflection dominates over the diffuse component of reflection $I_s(x, y) \gg I_d(x, y)$.

Figure 5.7 shows the measured transmitted irradiance in the sense regions 1 and 2 at 0° and 90° sensitive pixels for aluminum and plastic reflecting surface. The differences between the maximum and minimum transmitted irradiances for plastic and aluminum surfaces in both regions 1 and 2 are shown in table 5.1.

The range of the transmitted irradiance variation with incident polarization angle for plastic in sense region 1 is 190% and 340% higher than that of aluminium. In sense region 2 the range of the oscillations for plastic is 190% and 885% higher

Table 5.1 Transmitted radiance for plastic and aluminum

	Region 1		Region 2	
	Plastic	Alum.	Plastic	Alum.
$I_{max} - I_{min}(0°)$	0.66	0.15	0.466	0.16
$I_{max} - I_{min}(90°)$	0.67	0.23	0.985	0.10

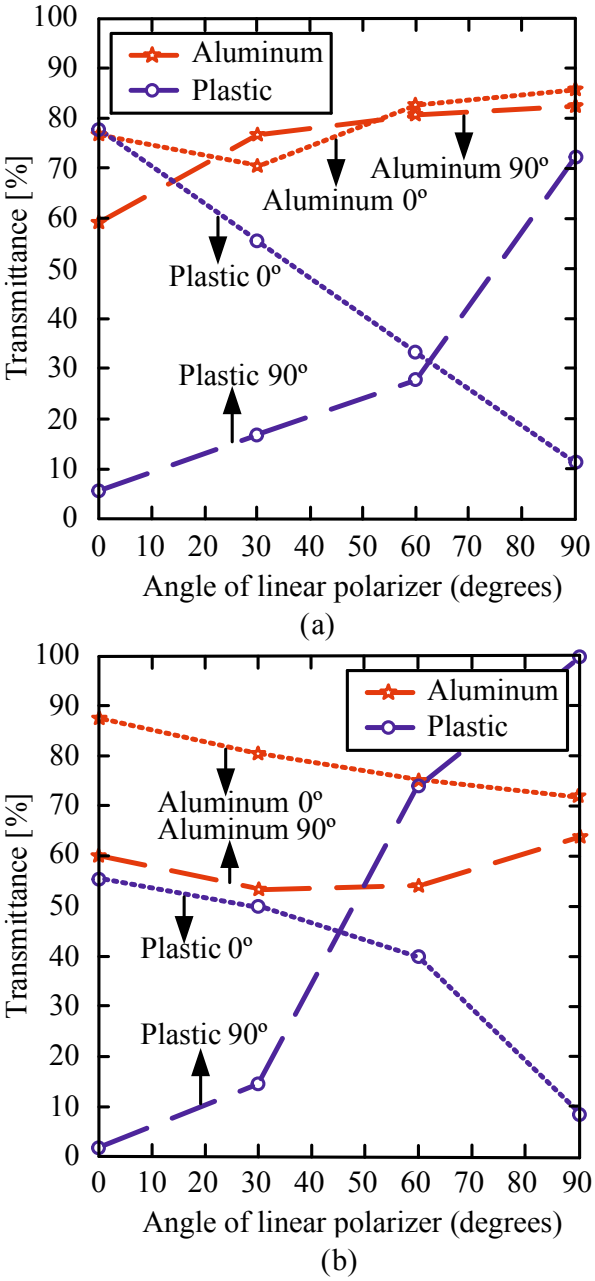


Fig. 5.7 Transmitted intensity at 0 and 90 polarization sensitive pixel in (a) sense region 1 (b) and sense region 2

than that of aluminum, for 90°and 0°polarization sensitive pixels respectively [9]. The transmitted irradiance oscillations for plastic at both polarization sense regions 1 and 2 are found to be much higher than those for aluminum. The differences in the transmitted irradiance are due to the difference in the reflection pattern of the light from the aluminum and plastic surfaces.

5.3.3 *Material Classification Using the Degree of Polarization*

The light reflected from the material surface is partially polarized. The polarization state of the reflected light can thus be represented as a sum of a completely polarized component and a completely unpolarized component [5], [7]. The degree of polarization is the ratio of intensity of the perfectly polarized light reflected to the total intensity of the reflected light. The difference between the maximum and the minimum transmitted intensities given by equations (5.22) and (5.23) shows the amount of reflected light that is completely polarized. The minimum transmitted radiance I_{min} is one half of the magnitude of the unpolarized light reflected from the object surface. If $I_p(x,y)$ is the partial polarization component and $I_{all}(x,y)$ is the total reflected component then:

$$\begin{aligned} I_p(x,y) &= {T_p}^{-1} A(x,y) = {T_p}^{-1}[I_x(x,y) - I_{min}(x,y)] \\ I_{all}(x,y) &= I_p(x,y) + {T_n}^{-1} B(x,y) \\ &= {T_p}^{-1}[I_x(x,y) - I_{min}(x,y)] + 2{T_p}^{-1} I_{min}(x,y) \\ &= {T_p}^{-1}[I_x(x,y) + I_{min}(x,y)] \end{aligned} \tag{5.29}$$

And the degree of polarization is then obtained from equation (5.29):

$$DOP = \frac{I_p(x,y)}{I_{all}(x,y)} = \frac{[I_{max}(x,y)] - [I_{min}(x,y)]}{[I_{max}(x,y)] + [I_{min}(x,y)]} \tag{5.30}$$

Equation (5.30) also indicates the portion of the reflected light which is completely polarized to the total amount of reflected light, denoting the partial polarization [5]. Equation (5.30) has a maximum value of one and a minimum value of zero. At a value of zero the reflected light is completely unpolarized, thus the diffuse component of the reflection dominates over the specular component. At the maximum value of one, the reflected light is completely polarized, thus the specular component of the reflection dominates over the diffuse component.

The transmitted intensities are obtained using the measurement setup shown in figure 5.5 as explained in section 5.3.2. The degree of polarization obtained for the polarization sense regions 1 and 2 is shown in figure 5.8. In figure 5.8, a *DOP* of 1 indicates the complete polarization of the reflected light and a *DOP* of 0 indicates the completely unpolarized state of the reflected light.

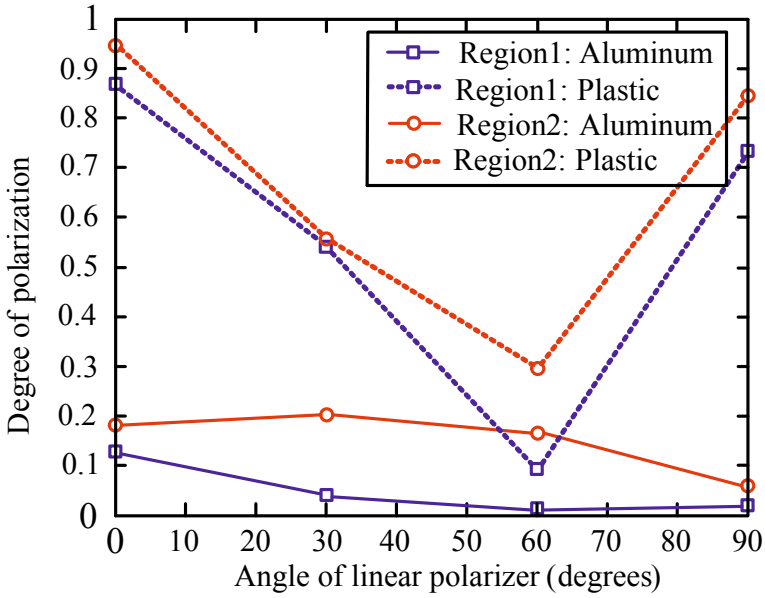


Fig. 5.8 Degree of polarization in sense regions 1 and 2

The maximum and minimum values of the *DOP* for the two reflecting surfaces of plastic and aluminum in the two polarization sense regions 1 and 2 are given in table 5.2.

Table 5.2 Degree of polarization for plastic and aluminum

	Region 1		Region 2	
	Plastic	Alum.	Plastic	Alum.
DOP (Max)	0.867	0.128	0.945	0.202
DOP (Min)	0.09	0.02	0.298	0.06

The degree of polarization is higher for plastic than for aluminum. It is observed from figure 5.8 that for plastic the maximum *DOP* is near one in both polarization sense regions, while for aluminum the maximum *DOP* in both polarization sense regions is less than 0.2. A higher *DOP* indicates higher amount of the reflected light being polarized, as stated in section 5.2.2.

It is further observed from figure 5.8 that as the specular angle of incidence is increased; the *DOP* tends to decrease for both plastic and aluminum. However after a certain specular angle of incidence, the *DOP* of aluminum continues to fall but the *DOP* of the plastic shows a sharp rise. From equation (5.17) it is seen that R_s never vanishes but R_p becomes zero when:

$$\tan(\varphi + \eta) = \infty$$
$$(\varphi + \eta) = \frac{\pi}{2} \tag{5.31}$$

From equation (5.31)

$$\sin \eta = \sin(\frac{\pi}{2} - \varphi)$$
$$\tan \varphi = n \tag{5.32}$$

where n is the index of refraction. When the angle of incidence satisfies the condition in equation (5.32), the electric field amplitude in the reflected wave has no component which lies in the plane of incidence, meaning that the entire component R_p gets refracted. The only component in the reflected wave is the one that is perpendicular to the plane of incidence. The reflected wave is completely polarized. The angle of incidence φ for which the reflected wave is completely polarized is known as the Brewster angle. The Brewster angle for plastics is near 60° [10], which explains the sharp rise of the *DOP* of plastic around 60°. For light absorbing material like metals R_p is never 0 and thus no such sharp increase in the *DOP* of aluminum is observed.

5.3.4 Material Classification Using the Stokes Parameters

As described in chapter 4, the polarization state of an electromagnetic wave can conveniently be described by a set of Stokes parameters. To our best knowledge, there is no reference available on using Stokes parameters to classify materials. Here the degree of polarization and linear degree of polarization obtained from the Stokes vector to classify materials into metals and dielectrics are used.

The measurement setup is the same as shown in figure 5.5. The outputs of the pixels sensitive to 90° and 0° are averaged over 30 frames. Simultaneously, the output of the intensity sensitive pixels with no metal grid is also recorded for varying transmission axis of the external linear polarizer. The pixel output of the 0° and 90° polarization sensitive pixels are then normalized with the output of the intensity sensitive pixels. The normalized outputs correspond to E_{x0} and E_{y0} in equation (4.15). The degree of linear polarization is calculated using equation (4.22). The Stokes parameters S_0 and S_1 can be obtained using equation (4.15). The Stokes degree of linear polarization in the regions 1 and 2 are shown in figure 5.9, where a 1 corresponds to a completely polarized state and a 0 indicates completely unpolarized state of the reflected light.

The Stokes degree of linear polarization for plastic has a maximum value of 1 and is higher than that for aluminum. It is further observed that the Stokes degree of linear polarization for plastics steadily decreases until it reaches the Brewster angle where the degree of polarization again increases to its maximum value. The maximum Stokes degree of linear polarization for aluminum obtained in the sense regions 1 and 2 are 0.25 and 0.38 respectively.

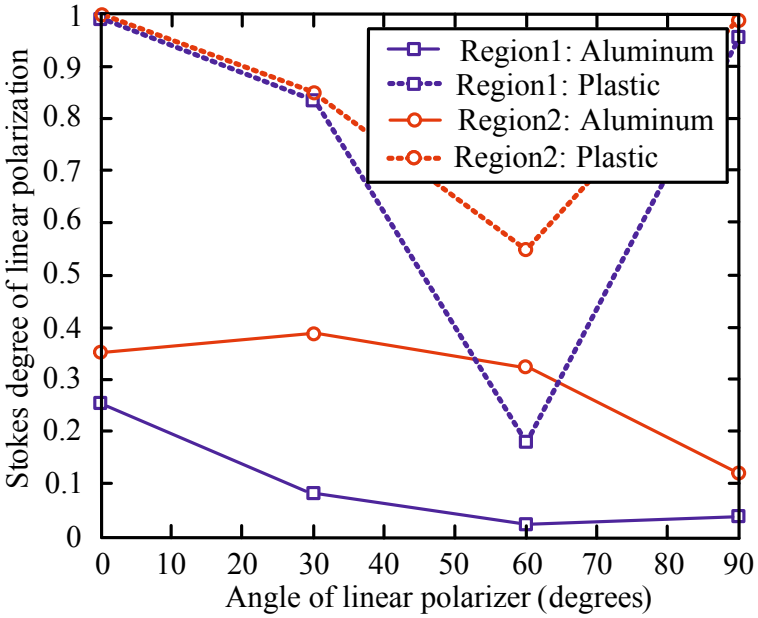


Fig. 5.9 Stokes parameters, degree of linear polarization in sense region 1 and 2

The behavior of the degree of polarization obtained using Stokes parameters (figure 5.9) are found to be very similar to the degree of polarization obtained using the maximum and minimum transmitted intensities (figure 5.8). It should indeed be the case as in linear polarization; the circular and elliptical forms of polarization are nearly absent and thus only the differences in maximum and minimum transmitted intensities are prevalent. In figure 5.8, aluminum has a degree of polarization less than 0.2 in both regions for varying angle of linear polarizer, while in figure 5.9 the degree of polarization is less than 0.4 in both regions. While for plastic the *DOP* is near one in both regions which steadily decreases till near Brewster angle and then further increases.

The above discussion was related to the degree of linear polarization obtained from the Stokes parameters. The Stokes degree of polarization is given by equation (4.21). The fourth Stokes parameter S_3 is usually ignored for natural light [11], since the phase information between orthogonally polarized light is difficult to calculate for natural light. Thus equation (4.21) can be modified with $S_3 = 0$ as

$$DOP = \frac{\sqrt{{S_1}^2 + {S_2}^2}}{S_0} \tag{5.33}$$

S_0 and S_1 are the same as for the linear degree of polarization. S_2 is calculated from equations (4.15). The phase difference, δ, in equation (4.15), is set to the specular

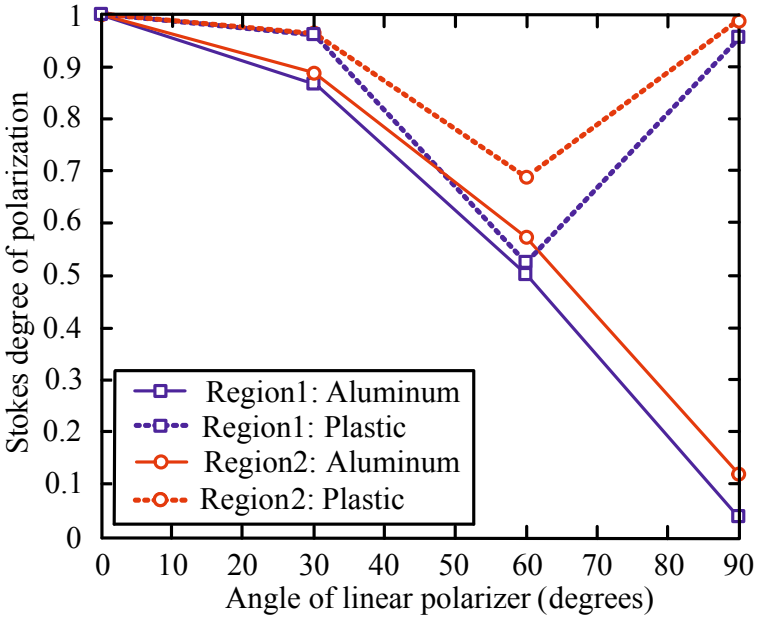


Fig. 5.10 Stokes degree of polarization in sense region 1 and 2

angle of incidence. The obtained degrees of polarization in the polarization sense regions 1 and 2 are shown in figure 5.10.

In figure 5.10, a Stokes *DOP* of 1 indicates the complete polarization of the reflected light while a 0 indicates the completely unpolarized state of the reflected light. Plastic shows a similar response to figure 5.9 while aluminum shows a steady decrease from the maximum value to minimum with increasing angle of linear polarizer. As the angle of specular incidence approaches the Brewster angle, the *DOP* in case of plastics increases to its maximum value while no such behavior is observed for aluminum neither in sense region 1 nor in region 2.

A threshold can be applied to the Brewster angle to classify between metals and dielectric using the degree of polarization computed using the Stokes parameters. Near the Brewster angle the Stokes *DOP* for plastics will be very near to 1 while in the case of metals it will be near 0.

5.3.5 Material Classification Using Polarization Fresnel Ratio

Wolff in [5] introduces the polarization Fresnel ratio *(PFR)* based on Fresnel reflectance model as a metric tool to classify materials into metals and dielectrics. The theoretical Fresnel reflection coefficients for aluminum and plastic are shown in figure 5.11.

The Fresnel reflection coefficient for aluminum is near 1, while that for plastic varies over the entire span from 0 to 1 for different specular angles of incidence.

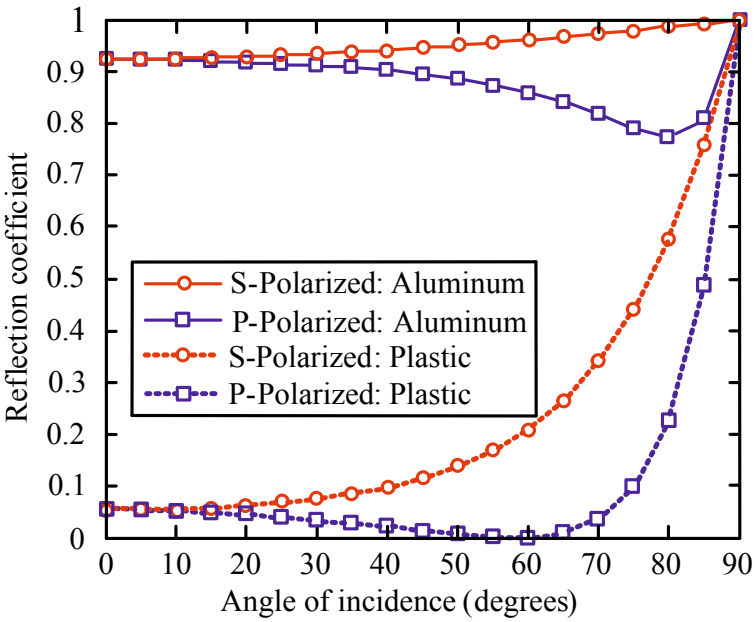


Fig. 5.11 Theoretical Fresnel reflection coefficients for aluminum and plastic at λ=500nm

The Fresnel reflection coefficient R_p is 0 near the Brewster angle for plastic and thus the *PFR* for dielectric (plastic) can become arbitrarily large, while the *PFR* for metals is limited.

A material with significant conductivity will have a significantly reduced *PFR* over a large range of specular angle of incidence. Since the conductivity of metals is higher than the conductivity of dielectrics, the *PFR* for metals is much smaller compared to that of dielectrics. As in metals the specular component of reflection is larger than the diffuse component of the reflection, the Fresnel coefficients satisfy $R_s(x,y) \approx R_p(x,y)$ [7]. Using this condition in equation (5.20) and (5.21) we get $A(x,y) \ll B(x,y)$ which from equation (5.22) and (5.23) means $I_{max}(x,y) \approx I_{min}(x,y)$, thus the *PFR* from equation (5.27) is nearly equal to 1 for all angles of incidence. The theoretical *PFR* for metal and dielectric surface is shown in the figure 5.12 [5].

There exists a threshold in the *PFR* values for metals and dielectrics. For the electromagnetic visible spectrum from 400nm to 700nm, it has been reported that the *PFR* for metals usually remains below 2 for most specular angle of incidence [5].

For varying transmission axis of the external linear polarizer, the pixel outputs averaged over 30 frames at the 0° and 90° polarization sensitive pixels are stored. The transmittance is computed by normalizing the output at 0° and 90° by the intensity at the intensity sensitive pixel in polarization sense region 1 and 2. The *PFR* is then calculated from the maximum $I_{max}(x,y)$ and minimum transmittances

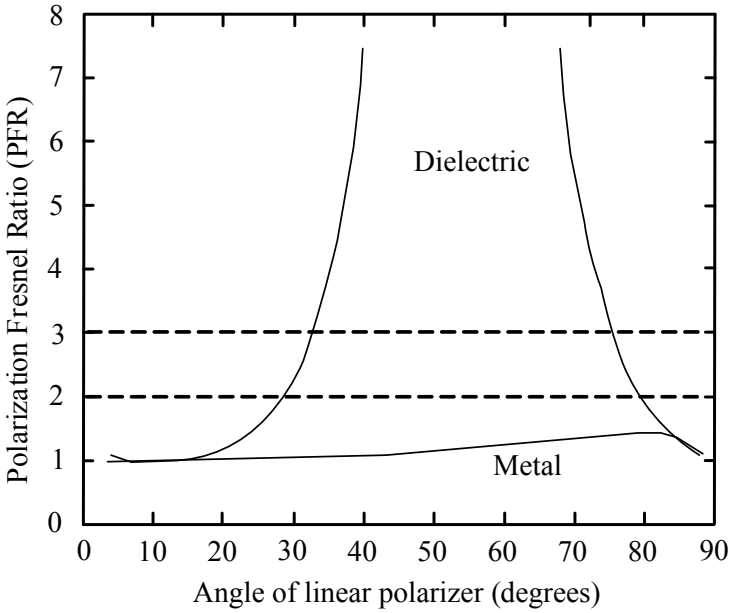


Fig. 5.12 Theoretical polarization Fresnel ratio [5]

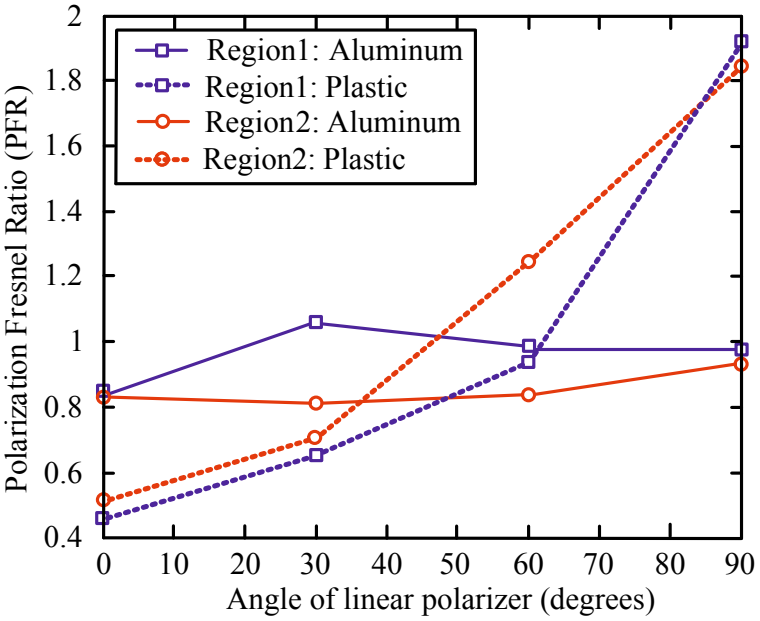


Fig. 5.13 Polarization Fresnel ratio in sense region 1 and 2

$I_{min}(x, y)$ determined using the same measurement setup shown in figure 5.5 and equation (5.27).

The experimentally obtained *PFRs* for aluminum and plastic in the polarization sense regions 1 and 2 are shown in figure 5.13.

The *PFR* for aluminum is found to be in the range of 0.8 to 1 for all specular angles of incidence while the *PFR* for dielectrics rapidly increases near the Brewster angle. Theoretically for a specular angle of incidence greater than the Brewster angle, the *PFR* for aluminum is always smaller than 2 [5]. A *PFR* value of nearly 2 can be considered to be a dielectric. The *PFR* computations using the polarization sense regions 1 and 2 offer a good match to the theoretical studies and thus can be used to classify the materials into metals and dielectrics.

5.4 Metal Classification

The Fresnel reflection theory can also be used to classify among conductive metallic surfaces. The Fresnel reflection coefficients depend on the index of refraction ϵ and the specular angle of incidence φ as shown in equations (5.17).

The basic equations for the propagation of a plane wave in a conducting medium differ from those relating to propagation in a transparent dielectric medium only in that the real constants are replaced by complex constants. The complex index of refraction ϵ is as denoted by equation (5.34).

$$\epsilon = n - i\kappa \tag{5.34}$$

where n is the simple index of refraction while κ is called coefficient of extinction. κ is a measure of how well a particular material scatters and absorbs electromagnetic waves. A material with low κ allows for easy transmission of the electromagnetic waves and vice versa. The index of refraction for dielectrics is a real number as the coefficient of extinction is negligible, while the index of refraction for metals is a complex number. The components of the index of refraction n and κ are related to electromagnetic physical parameters of a material surface [12] as

$$n^2 = \frac{\upsilon\gamma c^2}{2}[1 + \sqrt{1 + (\frac{\lambda\sigma}{2\pi c\gamma})^2}] \tag{5.35}$$

$$\kappa^2 = \frac{\upsilon\gamma c^2}{2}[-1 + \sqrt{1 + (\frac{\lambda\sigma}{2\pi c\gamma})^2}] \tag{5.36}$$

where σ is the conductivity of the material surface, υ is the electrical permittivity, γ is the magnetic permeability, λ is the incident wavelength of light and c is the speed of light in vacuum respectively.

The reflection and absorption of light by metals is also influenced by the skin effect [13]. It influences the optical behavior of noble and other metals with

high electrical conductivity (such as Cu, Al). A perfect conductor is characterized by infinitely large conductivity. However this means that a perfect conductor would reflect the entire incident light and it would not allow the penetration of the electromagnetic wave to any depth. Thus the reflection coefficients will be high for highly conductive materials while it will be low for lowly conductive materials. To observe the different reflection behavior for different metals of varying conductivity, the measurement setup of figure 4.5 was used. The linear polarization filter was removed and unpolarized light from the DC source was incident on the metal surface. All measurements were done for a light wavelength of 550nm. The reflected light intensity reaching the image sensor was measured using a light meter. Figure 5.14 shows the measured reflected intensity from metal surfaces of varying conductivity for the polarization sense regions 1 and 2 respectively.

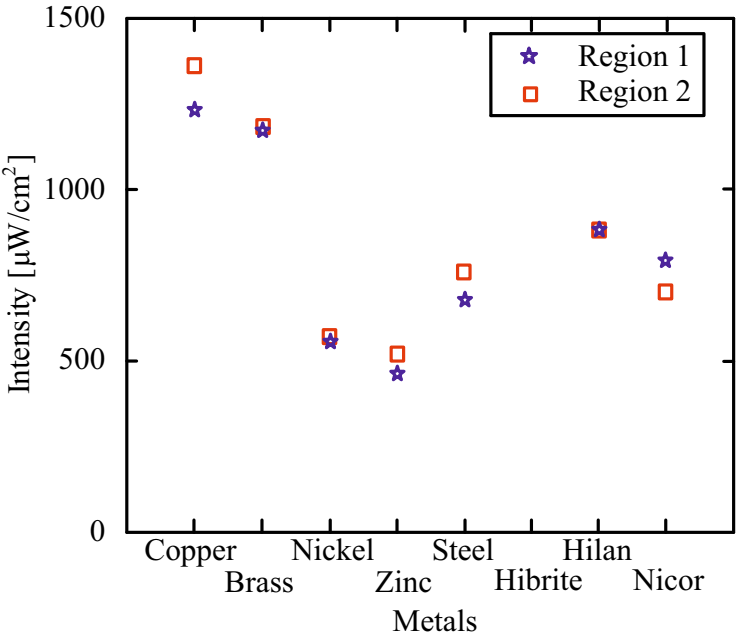


Fig. 5.14 Reflected intensity from various conducting metals in sense region 1 and 2

It is observed that copper and brass have a higher conductivity than the other metals used (nickel, zinc, steel, hilan, nicor and hibrite) and thus show a higher reflected intensity. This agrees with our previous discussion that higher conductive surfaces reflect the entire incident light due to skin effects. Hilan, nicor and hibrite, which are forms of steel, show relatively the same reflected intensity range. The conductivities of nickel and zinc are higher than that of steel, however the reflected intensity is lower. This could be due to the variations in the electrical and magnetic permeability for a given specular angle of incidence.

Table 5.3 Reflected intensity from various conducting metals in sense region 1 and 2

Metals	Intensity[$\mu W/cm^2$]	
	Region 1	Region 2
Copper	1360	1230
Brass	1184	1170
Nickel	575	557
Zinc	516	465
Steel	758	678
Nicor	703	791
Hibrite	796	691
Hilan	883	880

5.4.1 Metal Classification Using PFR

From equations (5.35) and (5.36) it can be inferred that, with all other physical parameters remaining constant, the index of refraction increases with an increase in the conductivity of the material. Thus it can be said that the Fresnel equations depend on the conductivity of the material surface. It was further observed in section 5.3.4 that as the conductivity increased from dielectrics to metals, the Fresnel reflection coefficients increased, reducing the *PFR* as a function of the specular angle of incidence. The variation in the conductivity among various metals will affect the Fresnel reflection coefficients which will further vary the *PFR*. Changes in the physical parameters in equations ((5.35) and (5.36) for various metal surfaces would result in different reflection behavior.

Table 5.4 shows the theoretical and experimentally obtained *PFR* for copper, zinc and aluminum in both polarization sense regions 1 and 2. The measurement set up is the same as the one shown in figure 5.5.

Table 5.4 Theoretical and experimental PFR values

Metals	Conductivity (/cm Ω)	Max. theoretical PFR	Experimental PFR	
			Region 1	Region 2
Copper	$0.596x10^6$	1.33	1.12	1.08
Alum.	$0.3776x10^6$	1.07	1.25	1.15
Zinc	$0.166x10^6$	1.31	1.21	1.28

The theoretical values are the maximum values where the *PFR* were averaged over different index of refraction corresponding to wavelengths in the visible spectrum [5]. The *PFR* for copper was found to be lower than that of Zinc and Aluminum in both polarization sense regions 1 and 2.

Among the selected three metallic surfaces, the conductivity of copper is the highest, and aluminum is more conductive than zinc. The higher conductivity of

copper produces higher Fresnel reflection coefficients which in turns results in lower *PFR*, seen in the experimentally obtained *PFR* in both polarization sense region 1 and 2. Zinc, being the least conductive, should have higher *PFR*, as seen in the measurements of polarization sense region 2. In polarization sense region 1 zinc shows lower *PFR* than aluminum. This could be the result of the variations in the specular angle of incidence of the light source, as the specular angle of incidence of the light source was not well controlled. However a clear distinction between the degrees of polarization for copper, zinc and aluminum is visible, and the *PFR* values are lower than 2 as stated in [5].

To extend the study of changes in the *PFR* with conductivity, more metallic surfaces were selected and the experimentally obtained *PFR* are shown in the figure 5.15.

The lower conducting materials steel, hibrite, hilan and nicor show higher *PFR* values in both the polarization sense regions 1 and 2, while highly conductive metals show lower *PFR*. The high *PFR* for the low conducting materials is due to reduced reflection and lower Fresnel reflection coefficients. A clear distinction can be observed among the *PFR* values of various test metal surfaces of varying conductivity.

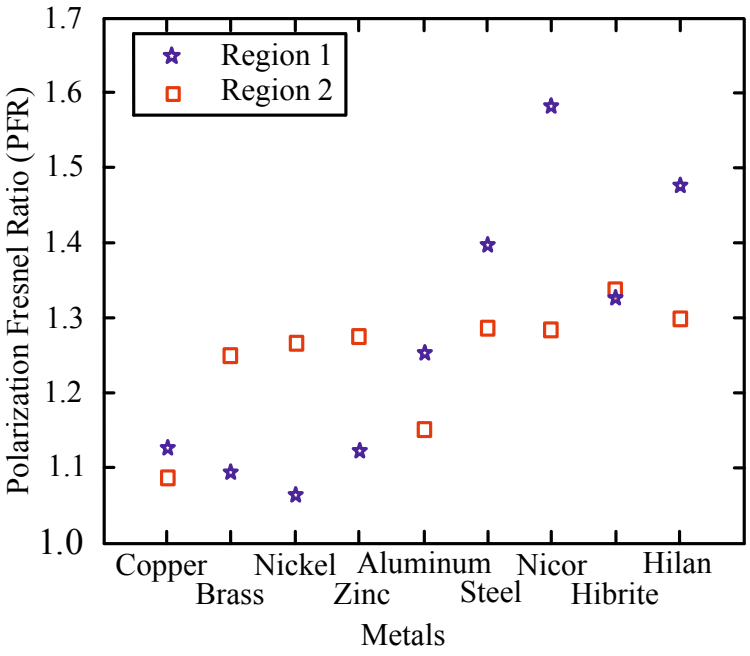


Fig. 5.15 Measured PFR for various conducting metals in sense region 1 and 2

Table 5.5 Measured PFR for various conducting metals in sense region 1 and 2

Metals	Experimental DOP	
	Region 1	Region 2
Copper	0.059	0.040
Brass	0.044	0.111
Nickel	0.030	0.117
Zinc	0.057	0.121
Alum.	0.112	0.070
Steel	0.165	0.125
Nicor	0.140	0.144
Hibrite	0.192	0.129
Hilan	0.226	0.124

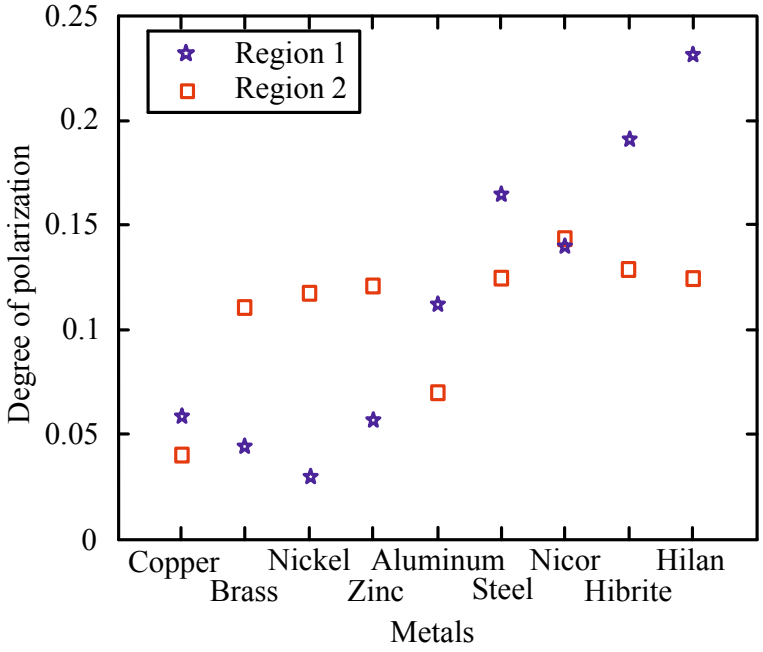


Fig. 5.16 Measured degree of polarization in sense region 1 and 2

5.4.2 Metal Classification Using Degree of Polarization

The *DOP* of highly conductive materials is lower than the *DOP* of low conducting materials. The partial polarization or *DOP* introduced in section 5.3.3 was also calculated for various metallic surfaces and the obtained *DOP* in the polarization sense regions 1 and 2 is shown in figure 5.16. The resultant plots show similar behavior as the *PFR* plots.

Table 5.6 Measured degree of polarization in sense region 1 and 2

Metals	Experimental DOP	
	Region 1	Region 2
Copper	0.059	0.040
Brass	0.044	0.111
Nickel	0.030	0.117
Zinc	0.057	0.121
Alum.	0.112	0.070
Steel	0.165	0.125
Nicor	0.140	0.144
Hibrite	0.192	0.129
Hilan	0.226	0.124

It is observed that the *DOP* of metal surfaces varies between 0 and 0.25 in both the polarization sense regions 1 and 2. The lower conducting metal surfaces, such as steel and its varieties occupy the higher band of the range while the lower band is occupied by highly conducting metal surfaces.

There is however a difference in behavior in the two polarization sense regions for both the *PFR* and *DOP* measurements, as seen in figures 5.15 and 5.16. In region 2, the difference between the copper and other conducting materials is more pronounced than in region 1. This can be explained by considering the presence of an additional 45° linear polarizer in region 2 along with the fact that the Fresnel coefficients measured are truly a regional average. For highly reflective surfaces the response of the 45° sensitive pixels increases the Fresnel coefficients average, thus further decreasing the *PFR* and *DOP*. For copper $R_s(x, y) \approx R_p(x, y)$, while for other conducting materials the average $R_s(x, y)$ increases due to the 45° polarizer response compared to $R_p(x, y)$, thus increasing the *PFR* and *DOP*. For very low conducting surfaces the 45° sensitive pixels do not affect the Fresnel coefficient measurements thus the measured average $R_s(x, y)$ decreases reducing the *PFR* and *DOP*.

5.5 Summary

- Materials are classified into metals and dielectrics. Metals are highly conductive, opaque and very reflective while dielectrics are less conductive and have low reflectivity.
- In dichromatic classification, materials are classified into optically homogenous and optically inhomogeneous. For metals the reflections are a function of the incident wavelength, while for dielectrics the reflections are independent of the incident wavelength.
- An unpolarized light can be polarized upon specular reflection. The reflected polarized light depends on the angle of incidence and the surface of the material.

It is thus possible to classify among reflecting surfaces based on the polarization of the reflected light.

- A light ray striking a surface will have a component which is reflected and a component which is transmitted or absorbed.
- The reflection occurring at the surface of the planar surface has two components: diffuse reflection and specular reflection. The specular reflection is a mirror like reflection while diffuse reflection is from internal scattering.
- For the reflected light, the intensity of the polarization component perpendicular to the specular plane is larger than the intensity of the polarization component parallel to the specular plane.
- The polarization state for the diffuse and specular components of the reflection depends on the reflecting surface. The measurement of the state of polarization of the reflected light serves as an indicator for the type of material surface.
- The magnitude of oscillations of the maximum and minimum transmitted irradiance due to the variation in the reflection pattern of metal and dielectric surface are different.
- The Polarization Fresnel Ratio *(PFR)* is the ratio of the perpendicular Fresnel coefficient to the parallel Fresnel coefficient. The Fresnel coefficients are derived from the Fresnel reflectance model.
- A material with significant conductivity will have a significantly reduced *PFR* over a large range of specular angle of incidence.
- The magnitude of variations of the reflected irradiance of the light after reflection is larger for dielectrics than for metals.
- The *PFR* for metals does not vary much with the variations in the linear polarizer angle and thus metals are easily distinguishable from dielectrics which show marked variance.
- The degree of polarization is the ratio of intensity of the perfectly polarized light reflected to the total intensity of the reflected light. The difference between the maximum and the minimum transmitted intensities are used to calculate the degree of polarization.
- If the maximum value of the degree of polarization is one, the reflected light is completely polarized.
- The specular component of the reflection dominates over the diffuse component.
- The degree of linear polarization obtained using Stokes parameters is also able to classify surfaces into metal and dielectrics.
- A perfect conductor reflects the entire incident light and does not allow any light penetration. Thus the reflection coefficients is high for highly conductive materials while it is low for lowly conductive materials.
- The polarization of the reflected component of the light wave varies with the conductivity of the metallic surface. This can be used to classify metallic surfaces as highly conductive and lowly conductive.
- The ability to classify materials using on-chip computation of degree of polarization *(DOP)* and the polarization Fresnel ratio *(PFR)* can be useful for real time material classification.

References

[1] Klinker, G., Shafer, S., Kanade, T.: Using a color reflection model to separate highlights from object color. In: Proceedings of International Conference on Computer Vision, pp. 145–150 (1987)
[2] Shafer, S.: Using color to separate reflection components. Color Research Applications 10, 210–218 (1985)
[3] Healey, G., Blanz, W.: Identifying metal surfaces in color images. In: Proceedings of Conference in Optics, Electro-Optics, and Sensors (1988)
[4] Healey, G., Binford, T.: Predicting material classes. In: Proceedings of DARPA Image Understanding Workshop, pp. 1140–1146 (1988)
[5] Wolff, L.: Polarization based material classification from specular reflection. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 1059–1071 (1990)
[6] Chen, H., Wolff, L.: Polarization phase based method for material classification and object recognition in computer vision. International Journal of Computer Vision 28(1), 45–56 (1999)
[7] Tominaga, S., Kimachi, A.: Polarization imaging for material classification. Optical Engineering 47(12), 123201-1–123201-14 (2008)
[8] Drude, P.: The theory of optics, p. 278. Dover publications Inc., New York (1959)
[9] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.: Integrated polarization analyzing CMOS image sensor. In: Proceedings of IEEE International Symposium on Circuits and Systems, pp. 621–624 (2010)
[10] Brewster, D.: On the laws which regulate the polarization of light by reflection from transparent bodies. Hilosophical Transactions of the Royal Society of London 105, 128 (1815)
[11] Gruev, V., der Spiegel, J.V., Engheta, N.: Advances in integrated polarization imaging sensors. In: IEEE Life Science Systems and Applications Workshop, pp. 62–65 (2009)
[12] Siegal, R., Howell, J.R.: Thermal radiation heat transfer. McGraw-hill, New York (1981) ISBN: 1560328398
[13] Gilberd, P.: The anomalous skin effect and the optical properties of metals. Journal of Physics F: Metal Physics 12, 1845–1860 (1982)

Chapter 6
Navigation Using CMOS Polarization Sensor

The navigational strategies of insects using skylight polarization are interesting for applications in autonomous agent navigation because they rely on very little information for navigation. The skylight polarization pattern for navigation varies in a systematic fashion both in plane (*e*-vector) and degree of polarization, depending only on the direction of the observation point relative to the angular position of the sun. This is found to be very efficient and reliable for real time navigation.

In this chapter, a polarization navigation sensor using the Stokes parameters to determine the orientation and position is presented. Section 6.1 discusses the two most common navigation algorithms prevalent in most animals: egocentric and geocentric. The working principle of the implemented algorithm is based on egocentric navigation, predominant in insects. The celestial compass based on skylight polarization is presented in section 6.2. Section 6.3 discusses the navigation compass employed by insects and in section 6.4 some of the implemented models for autonomous agent navigation based on the insect's model are presented. In section 6.5 the proposed polarization based compass is presented. The variation in the degree of polarization with changes in the polarized light can be used as a compass and conversely from the degree of polarization the incoming light ray direction can be determined. The computation of ellipticity and azimuthal angles allows for on-chip position detection based on the angle of the incoming light ray with little complexity. This can further be used as a sun position detector based on the skylight polarization. The ability to compute on-chip or in real-time the positional information would result in highly miniaturized navigational sensors.

6.1 Introduction

There are two types of spatial representation created by navigation. One type is a cognitive map or survey representation of space and the other is a route representation [1], [2]. Cognitive mapping or survey representation of the space defines the Euclidean relations (straight line distance and direction) among relevant landmarks within a coordinate reference system centered on the environment.

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 185–214.
DOI: 10.1007/ 978-3-642-34901-0_6 © Springer-Verlag Berlin Heidelberg 2013

Route representation or spatial representation of navigation relies on learning the points in the route.

The role of vision in navigation and cognitive mapping has been extensively studied [3]. Based on the cognitive mapping, navigation can be broadly classified into geocentric and egocentric [3], [4]. Humans navigate using geocentric form of navigation. In geocentric navigation, one uses its cognitive map and its orientation with respect to geocentric coordinates in order to set a course towards a goal. In this case visual cues such as landmarks become very important to generate the cognitive maps. In the absence of information about the location of nearby objects in the environment, humans have difficulty monitoring their travel trajectories, even for short paths. Egocentric navigation, is known to be used by some insect species such as the desert ant *(Cataglyphis fortis)* which relies on path integration, where the movement cues of the navigator are continuously integrated [4]. Animals using egocentric navigation are able to track distance and direction in order to estimate their position even in the absence of visual landmarks.

The progress in the field of autonomous agent navigation has been slower than expected, especially after the initial excitement and rapid advances in the early days of the research [5]. It is interesting to consider why autonomous navigation is so difficult. A survey into the available algorithms for autonomous agent navigation reveals that most of them are written to make them understandable for the human operator, mostly employing Cartesian coordinates *(x,y,z)* to represent the location of the feature [6]. Current sensors employed for navigation applications, and especially robotic navigations, use a generalized algorithm of capturing two-dimension images using a standard CMOS or CCD camera and then processing those captured image sequences for vector calculations to determine the motion vector and the direction vector with reference to the intimal position. The use of image sequences to determine motion and direction vectors is more related to the geocentric form of navigation, where the emphasis is more on the process of creation of two-dimension or three-dimension maps to guide the autonomous agents. This has the advantage of offering more control over the autonomous agent but demands large computational resources and lacks the robustness required for real-time applications. Thus these methods may not be the best solution for artificial autonomous navigation agents.

Insects, on the other hand are able to solve complex navigation problems in real-time relying on egocentric forms of navigation. Thus egocentric forms of navigation is best suited for autonomous agents than the geocentric form. This ability of insects to navigate effortlessly in complex environments without stressing their nervous system has already been a subject of research in robotics, and the cheap computational strategies that these insects use to navigate have been modeled by various researchers [7], [8], [9].

Path integration is the basis of vector navigation in egocentric forms of navigation [10], and it requires knowledge of the direction of travel and the distance travelled. In this book, the focus is on determining the direction of travel and not on distance measurements. To determine the direction of travel a reference direction is needed. The direction is then always expressed relative to the reference. In conventional autonomous agent navigation algorithms, multiple images captured

by a conventional camera and complex image processing algorithms are used to determine the angle of travel [6], [11]. The implementation of these algorithms needs very complex digital logic, and the image processing requires high power consumption and high bus bandwidth [12]. This limits the design of miniature and low-power vision-based navigation systems.

Insects such as the Saharan desert ant *(Cataglyphis fortis)*, use the position of the sun to determine their direction of travel [13]. Since the movement of the sun is constant and equal to approximately 1° every 4 minutes, it serves as a very good directional reference. To determine the direction of travel, ants use the celestial compass either based on the direct sunlight (sun compass) or on the pattern of the polarized skylight (polarization compass) [14]. Santschi [15] was the first to show that ants can use the sun as a compass cue. Genera of ants like *Messor* and *Monomorium* uses sun compass. In *Cataglyphis fortis* the polarization compass dominates over sun compass [14].

To determine the position of the sun using direct sunlight, conventional analog sun position sensors can be used. These sensors measure the position of the sun by allowing the light from the sun to pass through a pin-hole array and illuminate a certain region of the imaging array [16] as shown in figure 6.1.

The position of the illuminated region is then used to compute the altitude and position of the sun with respect to the sensor. State of the art digital sun sensors such

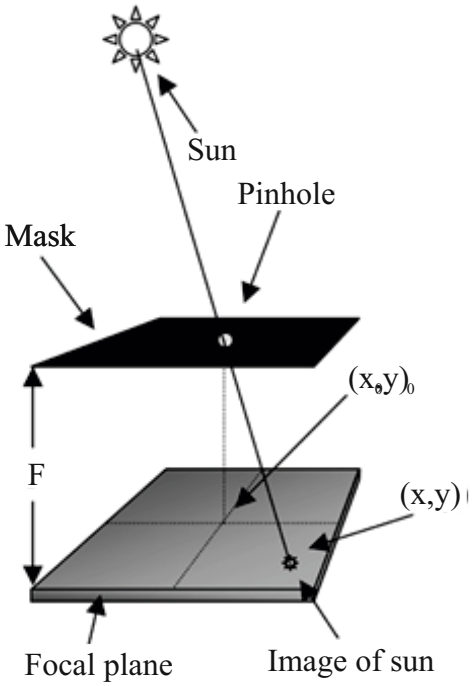


Fig. 6.1 Sun position detection model using image centroid

as [17] use the centroid method to compute the angular position of the sun while row profiling is used in the winner takes all *(WTA)* method proposed by [18]. The prerequisite of these sun sensors is that they need to see the sun, which is not always possible, for example in a cloudy day. Additionally, these sensors need an additional pin-hole array and digital processors to compute the centroid of the obtained image.

The other compass used by insects is the polarization compass. A polarized skylight navigation system uses the extensive pattern of polarized skylight generated by the scattering of the light rays on collision with the air molecules. Each partially polarized skylight ray exhibits a predominant vibration direction (e-vector) perpendicular to the plane of the scattering angle. The polarization pattern varies in a systematic fashion both in plane (e-vector) and degree of polarization, according to the position of the sun. These polarization patterns in the sky are used by ants as a reference for compass orientation. This makes their navigation pattern completely independent of external visual cues. The celestial compass based on the skylight polarization is described in section 6.2.

6.2 Celestial Compass Based on Skylight Polarization

Direct sunlight is unpolarized. When this unpolarized light enters the earth's atmosphere, it collides with the air molecules or is scattered because of the fluctuations in the air density. The scattering particles, the air molecules, are much smaller than the wavelength of the light striking them, thus the scattered intensity is the same in the forward and the backward directions. This scattering is often explained using the Rayleigh or molecular scattering theory, proposed in 1871 by Lord Rayleigh. The scattering caused by the randomly located air molecules is considered incoherent and thus can be assumed to be coming from only one particle. This scattering is dependent on the wavelength of the incoming light, where intensity of scattering is inversely proportional to the fourth power of the wavelength [19]. Blue light with wavelength at around 425nm scatters 5.5 times more energy than red light of wavelength 650nm. The higher scattering of the blue light explains the blue appearance of the sky.

The scattered light by the air molecules, or skylight, is partially polarized which means it is a composition of both the unpolarized natural light and also a linearly polarized component [20]. Circular and elliptical polarizations, explained in chapter 4, do not usually occur in the sky and thus only the linear polarization component is of interest. The scattered light has an e-vector oriented perpendicular to the plane of the scattering, i.e., perpendicular to the great circle passing through the sun and the point observed. Consequently, the e-vectors in the sky form concentric circles around the sky. This skylight polarization is mainly dependent on the angle between the viewing direction and the sun and the clearness of the sky in the viewing direction which affects the degree of polarization. There is an indirect relationship between the degree of polarization and the skylight intensity: for sun positions near the zenith, the degree of polarization is quite high while at the horizon the degree of polarization is weaker [21].

The degree of linear polarization is not constant over the entire sky but depends on the solar position of the sun as well as the atmospheric conditions. The strongest linear degree of polarization is observed during the sunrise and sunset at 90° from the sun position. The deeper and clearer the blue of the sky is, the stronger is the degree of polarization. However a total degree of polarization is never observed even when the viewing angle is 90° because secondary scattering and diffuse reflections always reduce the degree of polarization. Clear blue skies are strongly polarized while hazy skies exhibit very little polarization.

In clear sky, the polarization patterns are quite regular and depend very strongly on the position of the sun. The position of the sun in the sky relative to an observer on earth is defined by its solar elevation angle and its solar azimuth angle. The elevation angle is the angle between the line to the center of the sun and the horizontal plane. Horizontal plane is the reference plane for the solar elevation. when the sun is on the horizon, the elevation is 0° while when directly overhead the elevation angle is 90°. The azimuth angle is the angle between true south and the point on the horizon directly below the sun, the reference plane for the solar azimuth is the vertical plane running north-south through the poles. Since the *e*-vector orientation depends only on the plane of the scattering angle, the *e*-vector orientation is nearly independent of the atmospheric disturbance factors. The celestial *e*-vector pattern for two different elevations of the sun is shown in the figure 6.2.

The sun moves along its arc with uniform angular velocity of 15 degrees per hour, while the solar azimuth does not. Its angular velocity depends on the time of day, time of year and geographical latitude. Figure 6.2 shows two distinctive features. First, the gradient of the *e*-vector is related to the position of the sun and second, the pattern of polarization has mirror symmetry with respect to the plane defined by the solar meridian *(SM)* and the anti-solar meridian *(ASM)*. The solar meridian is the line between the sun and the zenith (highest point in the sky). The solar meridian rotates around the zenith as the position of the sun changes. Due to the westward movement of the sun, the *e*-vector pattern rotates around its zenith but retains the two important characteristics over the day, the mirror symmetry and that along the symmetry line the *e*-vectors are always perpendicular to the solar meridian [7].

A model to understand the polarization of the skylight is shown in figure 6.3. The transmission of light from the sun *(S)* to the scatterer *(P)* and the observer *(M)* is shown in the figure. The angle of elevation for the scatterer and the observer are ψ_s and ψ_p respectively.

The parallel and the perpendicular component of the polarized light radiates differently. The perpendicular component radiates omni-directional while the parallel component radiates proportional to $\cos(\psi_p - \psi_s)$ in figure 6.3, the difference of these two components produces the partial polarization of the skylight. The observer views the skylight through the linear polarizer with transmission axis at the angle ζ to the normal of the principal plane (x-axis). As the observer rotates the polarizer, he measures the maximum and minimum in the irradiance. This can be used to compute the degree of polarization [21] as:

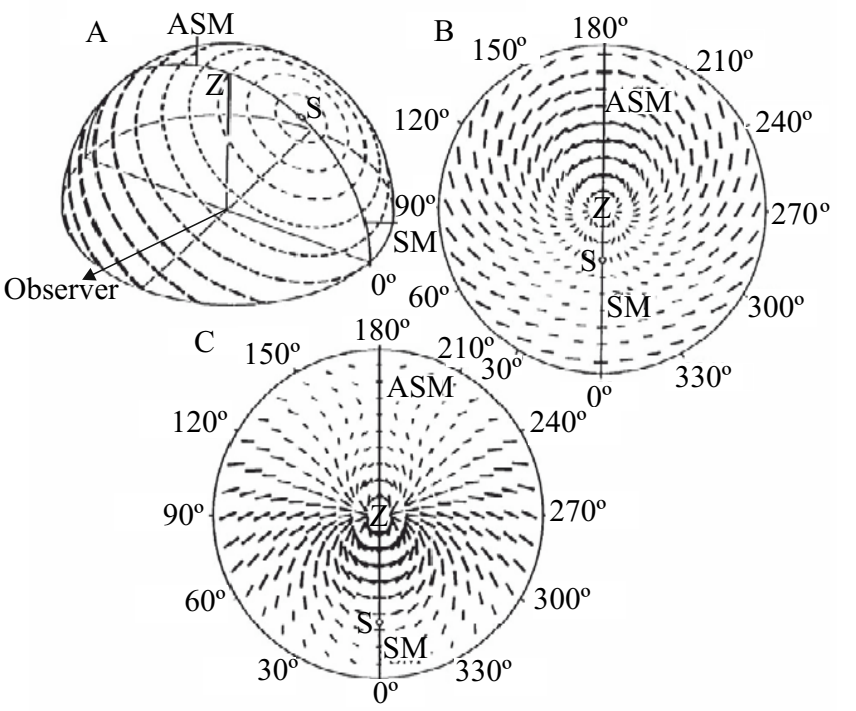


Fig. 6.2 Celestial e-vector pattern (A) Three-dimensional representation of the e-vector pattern, the orientation, length and width of each black bar corresponds to e-vector direction and the degree of polarization respectively. The symmetery line running through sun (S) and zenith (Z), is called "solar meridian" (SM) on the side of the sun and "anti-solar meridian" (ASM) on the opposite side. (B) Two-dimensional for sun elevation at 60° and (C) Two-dimensional for sun elevation at 24°.

$$DOP = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{\sin^2(\psi_p - \psi_s)}{1 + \cos^2(\psi_p - \psi_s)} \tag{6.1}$$

where *DOP* is the degree of polarization, I_{max} and I_{min} are the maximum and minimum transmitted intensity, ψ_s and ψ_p are the angle of the sun and the scatterer respectively. This scattered sunlight is partially polarized depending on the scattering angle, the angle between the incoming direct solar light rays and the outgoing skylight rays. The degree of partially linear polarized skylight can be expressed in terms of the scattering angle θ where $\theta = \psi_p - \psi_s$.

The degree of polarization as seen from equation (6.1) only depends on the difference in the angles $\psi_p - \psi_s$. Thus in the direction of the sun, when the scattering angle θ=0° ($\psi_p = \psi_s$), the skylight is completely unpolarized *(DOP=0)*. The maximum degree of polarization *(DOP=1)* would occur for a scattering angle of 90° ($\psi_p = \psi_s + \pi/2$).

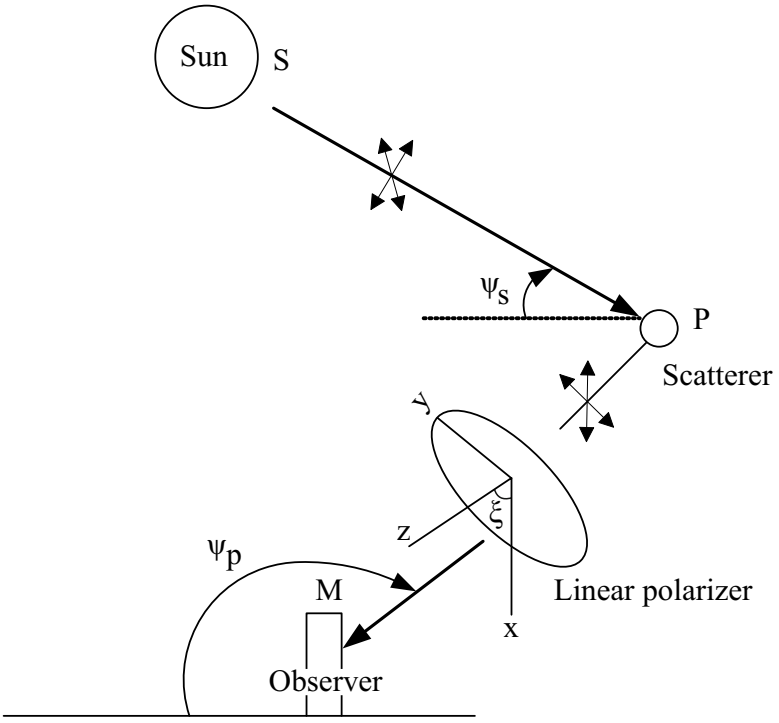


Fig. 6.3 Model to observe the skylight polarization

In figure 6.4, the theoretical degree of linear polarization predicted by equation (6.1) for varying sun elevation angle is plotted along with the measurement results found in literature. For the measurements the observer is viewing the the zenith sky $\psi_p = \pi/2$ as the sun rises (ψ_s increases) in figure 6.3. The two experiments conducted on Mauna Lao in Hawaii [20] and Bocaiuva in Brazil [22] both on high altitudes on a clear day, shows that the measured degree of polarization truly changes with the position of the elevation angle of the sun.

6.3 Navigation Using Polarized Light by Insects

Deserts are barren lands with almost no rain and vegetation. Navigation in such an environment is a challenge due to the absence of any visual landmarks. Humans who rely mostly on geocentric navigation depending on the visual cues and landmarks to form the cognitive map find it difficult to navigate in a desert even with a complex evolved brain. One habitant of such an environment is the Cataglyphis fortis *(Insecta: Hymenoptera: Formincidae)* shown in figure 6.5. For foraging activities such as search of food these ants travel up to twenty thousand times their own body length and go back home in a determined manner straightly without getting lost on the way [23]. This is truly amazing considering the absence of any visual cues or landmarks.

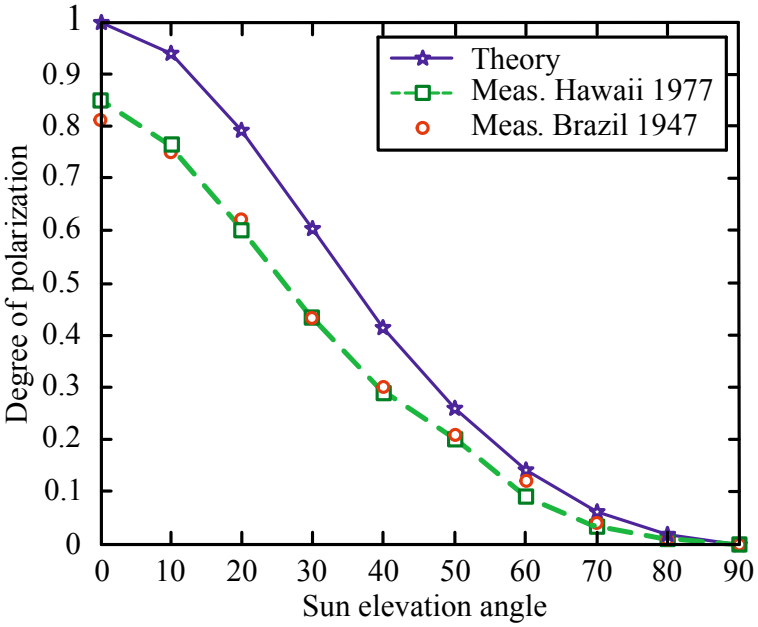


Fig. 6.4 Variation in degree of polarization with changes in the sun elevation angle

Fig. 6.5 Desert ant (Cataglyphis fortis)

Usually when foraging for food ants such as wood ants would mark the route of their travel by placing olfactory marks [24]. This serves two purposes, one to let other ants know the direction towards food and also helps in serving as a return path. This is not possible with Cataglyphis fortis as the sand in the desert is not cohesive

and would be carried away easily by wind. Also the dry environment and the hot surface temperature will evaporate the pheromones too quickly and thus cannot be used for stable marking. The navigation pattern of Cataglyphis fortis has perplexed biologists for a long time and theories about its navigational pattern have been postulated since 1904 when Pieron [25] argued that the ants while homing retrace the steps made during its outward journey by employing certain proprioceptive means. This argument was proven wrong by Cornetz in 1910 [26]. He proposed that the ants integrate their path rather than retrace the steps. This form of navigation is quite similar to that used by sailors at sea before the invention of global positioning system *(GPS)*. This concept of path integration by Cataglyphis fortis is now a widely accepted view [10], [27].

The path integration as the basis of vector navigation requires knowledge of the direction of travel and the distance traveled to compute the home vector. At each stage of foraging the ant continuously updates its home vector which points to the nest of the starting point of the travel. The compound eyes of ants are adapted to detect the polarization prevalent in the skylight [28]. There are three distinct functional regions in the compound eyes of Cataglyphis fortis: the dorsal rim area *(DRA)*, the dorsal area *(DA)* and the ventral area *(VA)* [29]. The microvilli of the *DRA* photoreceptors are aligned in parallel along the entire length of the cell from the distal tip of the rhabdom down to its proximal end and are perpendicular to other microvilli of the other photoreceptor cells [30]. These mutually orthogonal photoreceptors in the microvilli helps to obtain polarization antagonism. The polarization-antagonism has two important effects: it enhances *e*-vector contrast, which allows the neurons to respond to very low degree of polarization and it makes the system insensitive to the variations of absolute light level. That is, the neurons act as differential polarization detectors.

The *e*-vector patterns in the sky are not visible to humans, but ants use these vector directions as a reference for compass orientation. The *e*-vector information collected by the photoreceptors of the *DRA* is processed by polarization sensitive neurons in the optic lobe (*POL* neurons). Although most behavioral studies on the polarization compass are performed in bees and ants, due to the difficulty in recording from ant brains most electrophysiological data are obtained from orthopteran insects like the cricket [31]. The ants' *POL*-neurons exhibit opponency and wide visual fields, like the cricket. The *POL*-neuron of the cricket is shown in the figure 6.6.

The *POL*-neurons receive antagonist inputs from the receptor channels. The response function of the *POL*-neuron is a sinusoidal function of *e*-vector orientation with an excitatory and an inhibitory part and with the maxima and minima separated by 90°. The response function is the difference of the signals from the receptor channels and in principle determines the *e*-vector within the visual field of the neurons.

6.4 Navigation Using Polarized Light for Autonomous Agents

The principle of using polarization antagonism prevalent in the compound eye of insects as a polarization based compass for autonomous agents has been

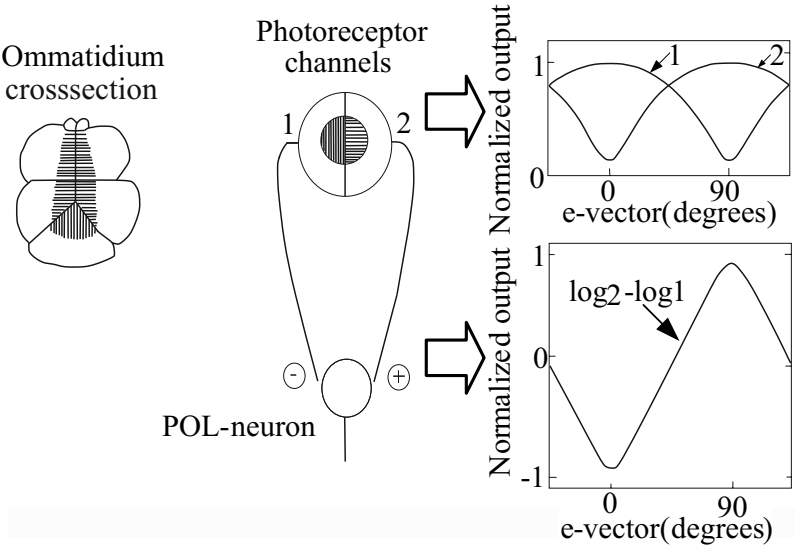


Fig. 6.6 (A) Ommatidum cross section of cricket. (B) Principle of operation of polarization opponent neurons (POL-neurons) [7].

successfully modeled by various researchers. Lambrinos et al. [7], [8] modeled the polarized light sensitive receptors (*POL*-neurons) in the Cataglyphis fortis using a pair of photodiodes with linear polarizers on top and a log ratio amplifier as shown in the figure 6.7. The pair of photodiodes with linear polarizers on top receives mutually orthogonal *e*-vector information, mimicking the receptors in the *DRA* region of the ommatidia of the compound eye of the insects. The response in each of the photodiodes is shifted by 90°. The signal from each diode is then fed to a log amplifier to obtain the logarithmic difference of the intensities received from the two photodiodes, similar in behavior as shown in figure 6.6.

Lambrinos [7] presented an analytical method of determining robot orientation based upon the polarized light neuron receptor geometry of the cricket. The three types of *POL*-neurons in a cricket are tuned to different *e*-vectors oriented approximately 10°, 60° and 130° relative to the length axis of the head. In the model proposed by Lambrinos et al., sensors with maximum sensitivity at the polarizer axis of 0°, 60° and 120° were placed on the robot and analytical expressions for orientation were derived based on models of sensor output. The sensor output was described by the following equation:

$$f(\varphi) = KI[1 + d\cos(2\varphi)] \tag{6.2}$$

where I is the total intensity, $I = I_{max} + I_{min}$, I_{max} is the maximum and I_{min} is the minimum intensity obtained from the photodiodes with linear polarizer in crossed-analyzer configuration (the two linear polarizers tuned to orthogonal *e*-vectors),

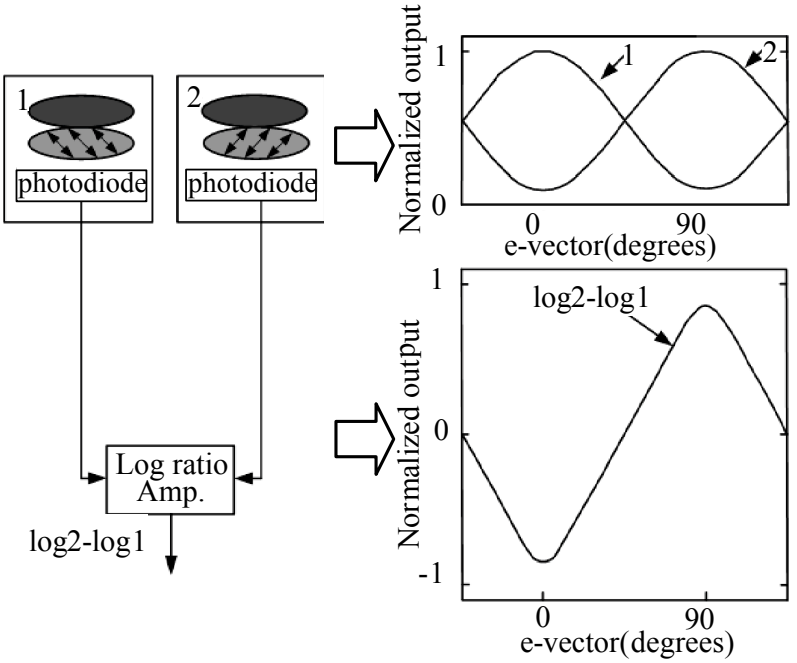


Fig. 6.7 Analog implementation of the POL-neuron [7]

d is the degree of polarization, φ is the orientation with respect to the solar meridian maximizing $f(\varphi)$, and K is a constant [32].

For the three *POL* units oriented at of 0°, 60° and 120° the output of the log amplifier can be expressed as,

$$f_1(\varphi) = \log\left(\frac{1 + d\cos(2\varphi)}{1 - d\cos(2\varphi)}\right) \tag{6.3}$$

$$f_2(\varphi) = \log\left(\frac{1 + d\cos(2\varphi - \frac{2\pi}{3})}{1 - d\cos(2\varphi - \frac{2\pi}{3})}\right) \tag{6.4}$$

$$f_3(\varphi) = \log\left(\frac{1 + d\cos(2\varphi - \frac{4\pi}{3})}{1 - d\cos(2\varphi - \frac{4\pi}{3})}\right) \tag{6.5}$$

where $f_1(\varphi)$, $f_2(\varphi)$ and $f_3(\varphi)$ are the outputs of the *POL* units oriented at 0°, 60°and 120°respectively.

To derive the compass information from equations (6.3), (6.4) and (6.5) either a scanning model or a simultaneous model [7] is used. In the scanning model the first task is to find the solar meridian and use it as a reference direction (0°). The sky is scanned using the system by rotating around its vertical axis to find the maximum.

The maximum value of the output signal is obtained when of the polarization filter is aligned to the solar meridian. After finding the reference direction, the information about the current state of the body can be used to find the direction of motion. In the simultaneous model, a look-up table is used to relate the obtained output values to the orientation and direction of motion instead of scanning the sky for a reference. While scanning the sky the maximum output obtained is used as a reference.

Equation (6.2) is dependent on the degree of polarization. The degree of polarization changes during the day depending on the position of the sun as shown in figure 6.2. In the scanning model of the compass determination, the degree of polarization is naturally eliminated as only the maximum output value is of interest. In the simultaneous model the degree of polarization has to be continuously updated based on the time of day in the look-up table or by normalizing the outputs by first finding the anti-logarithm of the output equations and then applying a sigmoid function. A sigmoid function is a real valued differentiable function whose first derivative is bell shaped and has a single non-negative or non-positive local maximum.

Though Lambrinos et al. in their robots have successfully demonstrated compass detection capability, an important disadvantage of their system is the use of three cameras. Three cameras with 0°, 60° and 120° orientations have to be monitored which is expensive in terms of cost, area, complexity of the algorithm and also the computation power needed to implement the algorithms. Kane et al. [9] proposed a much simpler method where a normal camera with a linear polarization filter is used to take two images, the second image taken with the polarization filter set orthogonal with respect to its position in the first image. From these two images a mean intensity function as a function of polarization angle was derived similar to the form proposed in [7] for the photosensitive diodes.

To solve for the three unknowns in the equation (6.2) Kane et al. use three different images with the polarization transmission axis set to 0°, 45° and 90° respectively. Rearranging equation (6.2) and accounting for images from the different transmission axis, three different image equations can be written.

$$f_1(\varphi) = KI[1 + d\cos(2\varphi)] \tag{6.6}$$

$$f_2(\varphi) = KI[1 + d\cos\left(2\varphi - 2\frac{\pi}{4}\right)] \tag{6.7}$$

$$f_3(\varphi) = KI[1 + d\cos\left(2\varphi - 2\frac{\pi}{2}\right)] \tag{6.8}$$

The set of equations is solved at each pixel of the image to compute the orientation angle φ.

This model also suffers from the disadvantage of using an external linear polarizer which needs to be rotated to have three images with three different transmission axis of the linear polarizer.

Both the models of Lambrinos et al. and Kane et al. are also not very well compensated for the change in the degree of polarization during the day. The models also need an ephemeris compensation algorithm. The ephemeris function is a function that describes the change of the sun's azimuth over time and it depends on the season and the geographical latitude. It is known that the rate of change of the sun's azimuth is not constant: it is faster around noon and slower in the morning. Insects such as bees and ants also suffer from the inherent changes in the ephemeris function but are known to refine their compass detection ability with experience.

6.5 Polarization Based Compass

The models summarized in section 6.4 by Lambrinos et al. and Kane et al. are very similar. The former generates an output image by subtracting the individual image obtained by the two mutually orthogonal photodiodes while the latter generates an output image by adding the two mutually orthogonal images. In the designed image sensor described in this book, these functionalities can be easily incorporated because the individual pixels are polarization sensitive as shown in chapter 4.

6.5.1 Measurement Setup

The measurement setup is the same described in chapter 4. The DC light source generates unpolarized light which is polarized by a linear polarizer. The linear polarized light intensity is then sensed by the imager and the analog output of the imager after digitization using an external *ADC* is fed to the PC for processing. The corresponding analog outputs of the pixels sensitive to 0°, 45° and 90° in the polarization sense regions are used to compute the Stokes parameters.

For the first version of the sensor the Stokes parameters were computed off-chip to have a proof of concept. The circuits required for the computation of Stokes parameters and other digital processing for the implementation of the algorithm can be integrated on chip.

6.5.2 Measurement Results

To validate the functionality as described in section 6.4, the pixel intensities for the 0° and 90° polarization orientation in the two polarization sense regions are recorded. The normalized pixel intensities for 0° and 90° polarization sensitive pixels in sense regions 1 and 2 are shown in figure 6.8. For both polarization sense regions the 0° linear polarizer pixel has the maximum at 0° and the minimum at 90°. Similarly, for the 90° polarizer pixel the maximum is observed at 90° and the minimum at 0°.

To test the compatibility of the designed model with the model used by [7], the logarithmic difference of the mean intensities of the two linear polarizations was

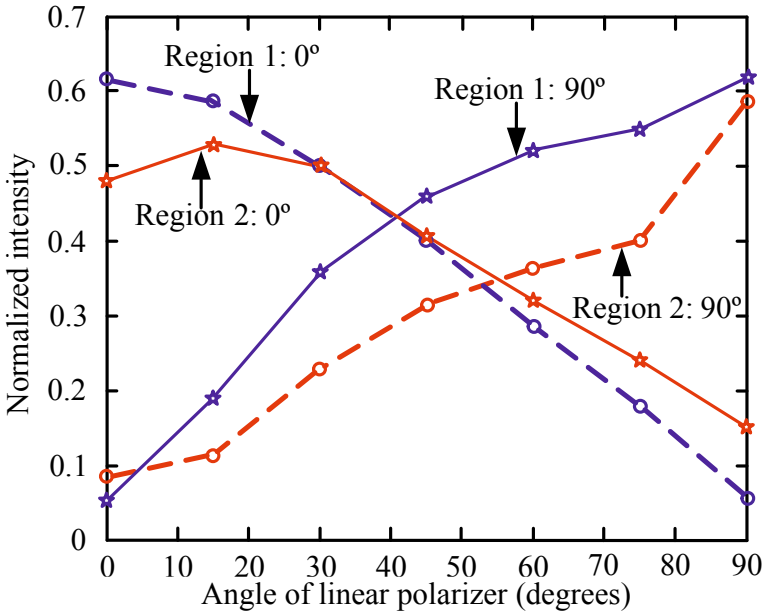


Fig. 6.8 Normalized intensity obtained for 0° and 90° polarization

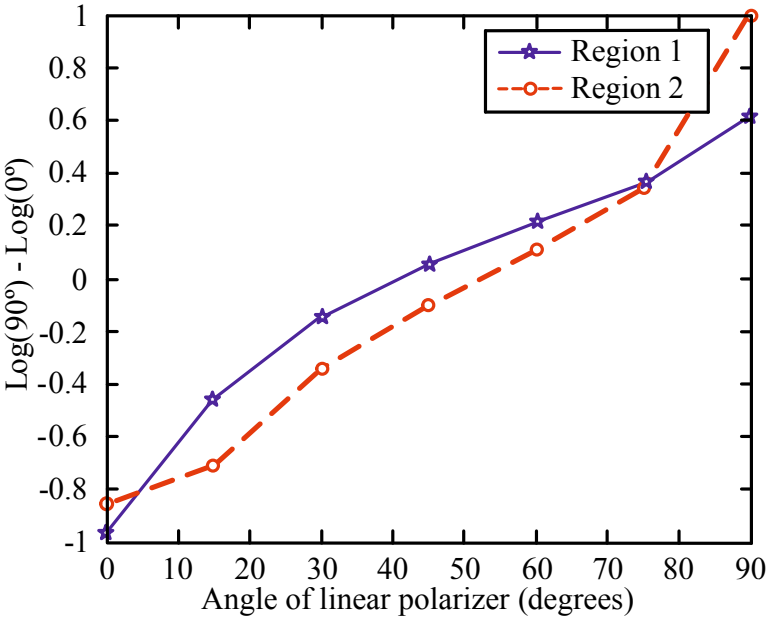


Fig. 6.9 The logarithmic difference of the orthogonal signals in sense region 1 and 2

computed off chip and is plotted in figure 6.9. A very similar characteristic as the one shown by [7] and [8] was achieved for both sense regions.

For both polarization sense regions, the logarithmic difference of the two orthogonal signals was found to have a minimum at 0° linear polarizer angle and a maximum at 90° linear polarizer angle. The logarithmic difference shows an approximately linear increase for increasing linear polarizer angle [33]. This variation in the logarithmic difference can be conceptually used to determine the angle of linear polarization, and based on a reference or look-up table it can be used as a compass.

A different approach is taken here and instead of focusing on polarization differential image or polarization summation image, the variation in the degree of polarization with respect to the orientation of the *e*-vector is studied. The image intensity from the photodiodes in the models described in section 6.4 is a function of the degree of polarization which changes over the course of the day. Lambrinos et al. offsets the effects of the variation in the degree of polarization by normalizing the output and Kane et al. did not compensate for it. The degree of polarization behavior with the orientation angle is evaluated and found that this degree of polarization information can be used to extract the orientation angle and hence can serve as a compass clue.

The variations in the degree of polarization in an image (φ) for changes in the orientation angle φ with respect to the solar meridian is given by equation (6.2). An inverse relationship between the degree of polarization d and the orientation angle φ is obtained.

$$d = \frac{1}{\cos(2\varphi)\left[1 - \frac{f(\varphi)}{KI}\right]} \tag{6.9}$$

Various methods of calculating the degree of polarization from the measured transmitted intensities through the implemented embedded metallic wire grid micro-polarizer have been discussed in chapter 4 and 5. The Stokes degree of linear polarization ((4.22)) is shown in figure 6.10. The degree of polarization calculated using the maximum and minimum transmittance, equation (5.26), is shown in figure 6.11. Both plots show an inverse relationship between *e*-vector orientation angle and the degree of polarization. The Stokes parameters S_1 and S_2 used to calculate the Stokes degree of linear polarization are obtained as:

$$\begin{aligned} S_0 &= {I_{0^\circ}}^2 + {I_{90^\circ}}^2 \\ S_1 &= {I_{0^\circ}}^2 - {I_{90^\circ}}^2 \end{aligned} \tag{6.10}$$

where I_{0° is the intensity of the light after passing through a horizontal linear polarizer, I_{90° is the intensity after a vertical linear polarizer. Figures 6.10 and 6.11 show that both the *DOLP* obtained either by the Stokes parameters or by the *DOP*

obtained from the maximum and minimum transmittance vary between +1 to -1 as the polarizer angle is varied from 0°to 90° [33].

The variations in the degree of linear polarization with respect to the orientation angle can be used as a compass. This approach has the advantage of using fewer computations than the model proposed by Lambrinos et al., where at first the system output was logarithmized and then anti-logarithamized and normalized to remove the dependence on the degree of polarization.

The *DOP* in its simplest form is represented by the Polarization Fresnel Ratio *(PFR)* described in section 5.3.5. The *PFR* is the ratio of the perpendicular Fresnel coefficient to the parallel Fresnel coefficient. The Fresnel coefficients describe the reflection and transmission coefficients of the light wave at an interface. The *PFR* can be roughly estimated as:

$$PFR = \frac{P_{per}}{P_{par}} \tag{6.11}$$

where P_{per} is the power of the perpendicular polarization and P_{par} is the power of the parallel polarization.

The pixel intensities obtained from the pixels covered with wire grid polarizer tilted by 0° and 90° in both sense regions are used to compute the *PFR*. The variation in the *PFR* with varying angle of linear polarizer is shown in figure 6.12.

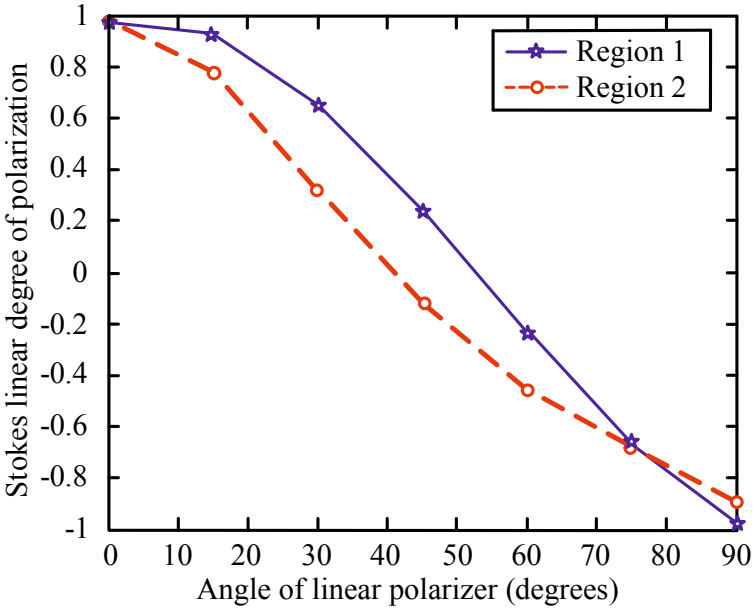


Fig. 6.10 Stokes linear degree of polarization for sense region 1 and 2

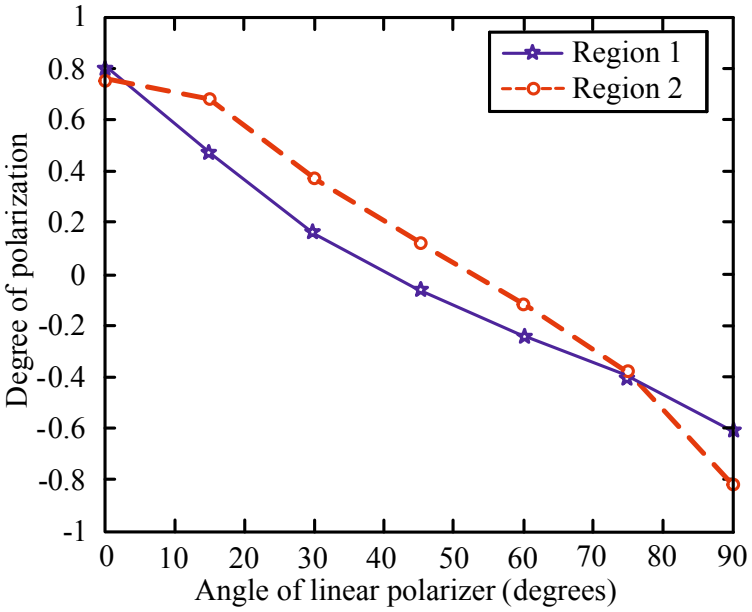


Fig. 6.11 Degree of polarization for sense region 1 and 2

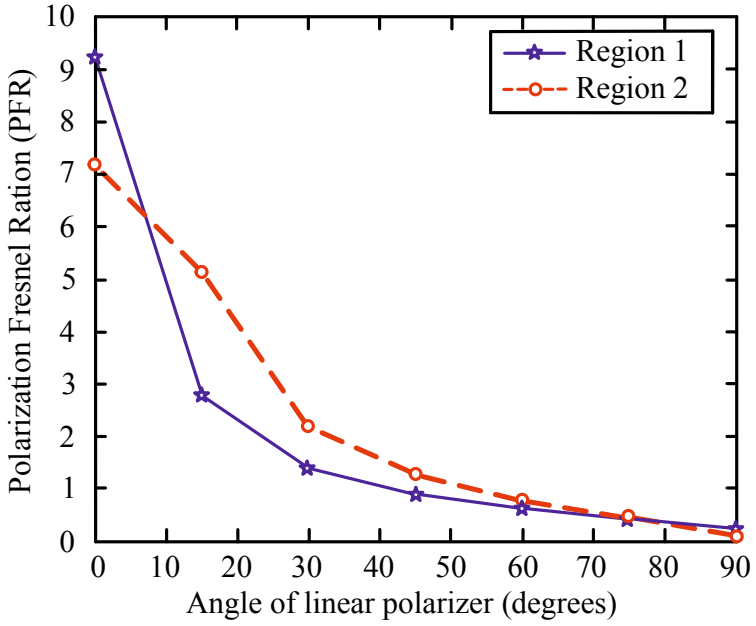


Fig. 6.12 Polarization Fresnel Ratio (PFR) for sense region 1 and 2

From figure 6.12, it is observed that the *PFR* has a maximum value when the polarizer angle is 0° and steadily decreases to a minimum when the polarizer angle is 90° [33]. The decrease is steeper for a lower angle of the linear polarizer than for the higher values. Although the variations in the *PFR* with varying polarizer angle are not linear, the variation for various angles of linear polarizer is still suitable to be used as compass to detect the direction of the incoming polarized light ray.

6.6 Incoming Light Ray Direction Detection and Sun Position Detection

In section 6.5 the variations of the degree of polarization with changes in the orientation angle were observed. In this section it is shown that it is also possible to compute the *e*-vector orientation angle from varying degree of polarization. A polarization navigation sensor principle which uses Stokes parameters to determine the incoming polarized light ray direction is described.

6.6.1 Measurement Setup

The experimental setup shown in figure 6.13 was built for indoor measurements using a DC light source.

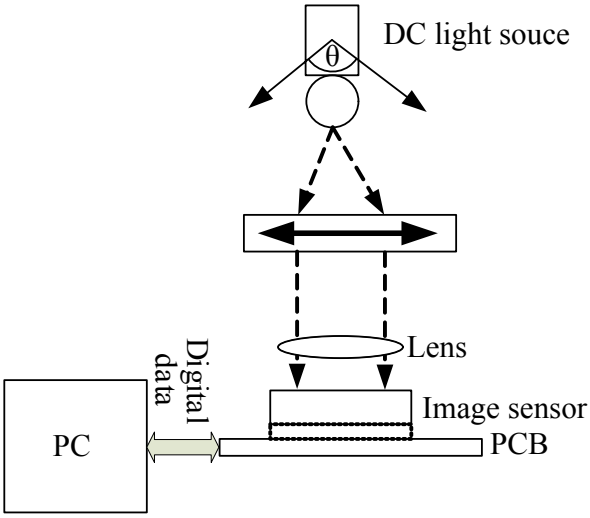


Fig. 6.13 Measurement setup for incoming light ray direction

The measurement setup is very similar to the one described in section 6.5.1, only here the DC light source is not fixed but it can be moved to have an angular direction. The sensor is illuminated with a polarized light obtained by passing the light from

a DC light source through an external linear polarizer. The corresponding analog outputs of the pixels sensitive to 0°, 45° and 90° in the polarization sense regions are recorded. These measurements are then sent to the PC for further processing. For the first version of the sensor the Stokes parameters were computed off-chip to have a proof of concept.

6.6.2 Measurement Results for Incoming Light Ray Directions

The Poincaré sphere described in chapter 4 is used to display the Stokes parameters and is characterized by an ellipticity angle and an azimuthal angle. The ellipticity angle and the azimuthal angle of the Poincar sphere, computed from the Stokes parameters are shown to be correlated with the incoming polarized light ray direction.

6.6.2.1 Incoming Light Ray Direction Detection Using Ellipticity Angle

To measure the ellipticity angle projected by the incoming light ray, the angular position of the DC light source in figure 5.13 was varied to change the incoming light ray direction. Using equations (4.21) and (4.24), the ellipticity angle can be expressed as

$$\chi = \frac{1}{2}\sin^{-1}\left(\frac{S_3}{\delta x S_0}\right) \tag{6.12}$$

where χ is the ellipticity angle, δ is the degree of polarization and S_0 and S_3 are the Stokes parameters. The variation in the measured ellipticity angle with the variation in the angular position of the DC light source is shown in figure 6.14.

The ellipticity angle (χ) as shown in equation (6.12) depends on the degree of polarization. The measured ellipticity angle for 10°, 15° and 30° incidence is shown as "Stokes *DOP*" in figure 6.14, where the degree of polarization is computed using the equation (4.21). The computation of the Stokes degree of polarization from the Stokes parameters needs squaring and square root arithmetic operations. These operations are relatively difficult to implement on-chip.

The calculated ellipticity angle using the partial polarization computed from equation (5.31) is shown as "Partial Polarization" in figure 6.14. This computation is easier to implement on chip as it would only need a differential amplifier to compute the difference and an analog divider. A real time implementation model of an analog divider is discussed in section 6.7. In the skylight polarization only the linear component of the polarization is dominant, and the circular and elliptical polarization components are usually absent. Thus for experiments under open sky, only the linear degree of polarization will be measured. The measurement of the ellipticity angle using the Stokes degree of linear polarization as in equation (4.22) is shown in figure 6.14 as "Linear *DOP*".

The correlation coefficient indicates the strength and direction of a linear relationship between two random variables and can be calculated using equation (6.13)

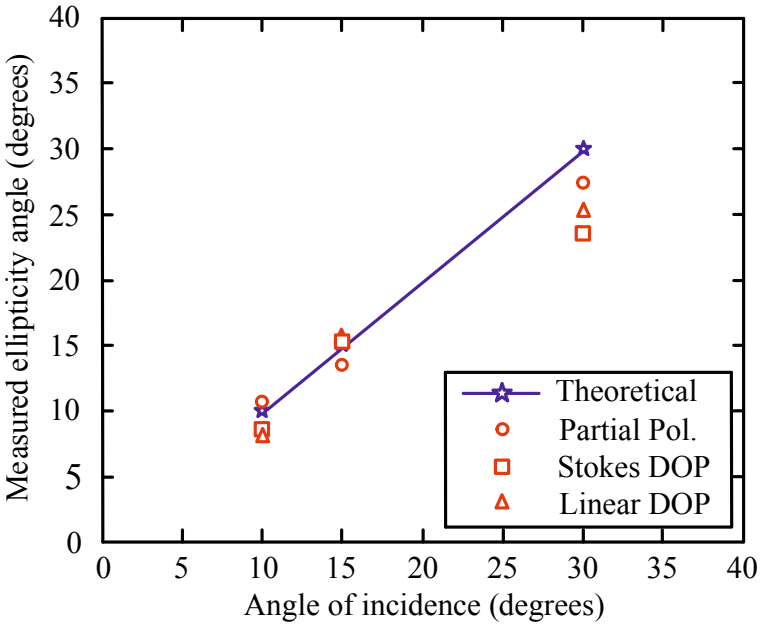


Fig. 6.14 Measurement of incident light ray angle

$$\text{correlation} = \frac{[N\sum xy - \sum x \sum y]}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \tag{6.13}$$

where N is the number of elements and x and y are the two variables.

Correlation coefficients of 0.98, 0.94 and 0.97 are obtained between the theoretical and measured results for Stokes *DOP*, Partial polarization and linear *DOP* respectively [34]. The high values of the correlation coefficients imply a strong correlation between the theoretical and the measured results. Thus the incoming polarized light ray direction can be related to the ellipticity angle of the Poincaré sphere.

6.6.2.2 Incoming Light Ray Direction Detection Using Azimuthal Angle

The azimuthal position (ψ) is computed from the Stokes parameters using equation (4.15) and (4.24). The theoretical and experimentally obtained azimuthal angle in the sense regions 1 and 2, or equivalently, the polarization angle measured while varying the transmission axis of the linear polarizer, is shown in the figure 6.15.

The correlation coefficients of 0.997 and 0.98 in the two regions indicate a strong correlation between the theoretical and the measured results. A linear fit error for the measured angle of linear polarization is computed to be 2.3% and 1.1% in the sense region 1 and 2 respectively [35]. For a similar configuration using an

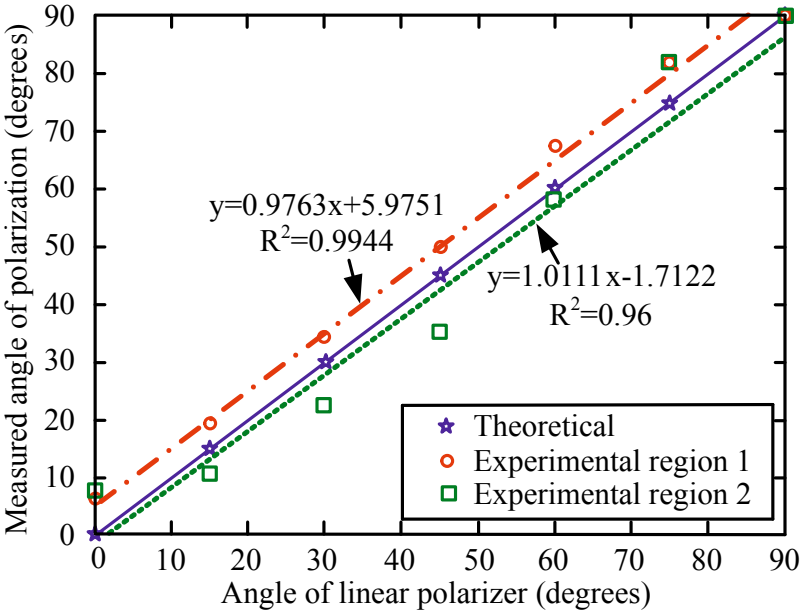


Fig. 6.15 Measured angle of polarization in polarization sense region 1 and 2 using equations (4.15)

organic micro-polarizer, the error is reported to be 2.2% [36] and an error of 1.6% is reported when aluminum nanowires is used with a wire grid pitch of 70nm [37] . The organic micro-polarizer exhibits a very high *ER*, thus the polarization intensity measurements are more accurate. The fabrication of these organic micro-polarizer's require very specialized processes, while the proposed method in this book uses standard CMOS technology processing steps and produces comparable results.

The polarization sense region 2 has an additional 45° linear polarization sensitive pixel (figure 4.32). The Stokes equations are then modified to use the 45° linear polarization information as:

$$\begin{aligned} S_0 &= {I_{0^\circ}}^2 + {I_{90^\circ}}^2 \\ S_1 &= {I_{0^\circ}}^2 - {I_{90^\circ}}^2 \\ S_2 &= {I_{0^\circ}}^2 - {I_{45^\circ}}^2 \end{aligned} \tag{6.14}$$

where I_{0°, I_{90° and I_{45° are the intensity of the light after passing through linear polarizer oriented at 0°, 90°and 45°. The fourth Stokes parameter S_3 is neglected as the light is assumed to be completely linearly polarized. The azimuthal position (ψ) is then computed using equations (4.24) and is shown in figure 6.16.

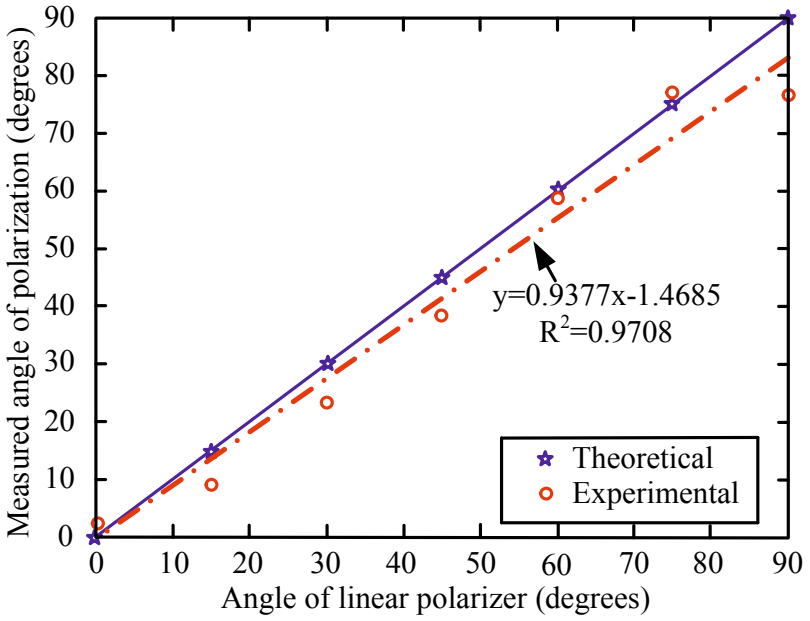


Fig. 6.16 Measured angle of polarization in polarization sense 2 using equations (6.14)

A linear fit error for the angle of linear polarization measurements is computed to be 3.07%. A correlation coefficient of 0.97 is obtained between the theoretical and the experimental values [35].

The Stokes parameters used to obtain the azimuthal position (ψ) in figure 6.15 and figure 6.16 require the computation of the square of the pixel intensities, which is relatively difficult for an on-chip computation. To simplify the computation the Stokes equations can be further simplified as shown in equation (6.15).

$$\begin{aligned} S_{m0} &= I_{0^\circ} + I_{90^\circ} \\ S_{m1} &= I_{0^\circ} - I_{90^\circ} \\ S_{m2} &= I_{0^\circ} - I_{45^\circ} \end{aligned} \tag{6.15}$$

where S_{m0}, S_{m1} and S_{m2} are modified Stokes parameters, computed from the intensities. Equation (4.24) is again used to compute the azimuthal position ψ and is plotted in the figure 6.17.

A linear fit error for the measured angle of linear polarization is computed to be 6%. A correlation coefficient of 0.985 is obtained between the theoretical and the experimental values showing strong correlation [35]. The angular positional information obtained by equations (4.15) and (6.14) is very similar to those obtained by equations (6.15) as shown in the figures 6.16 and 6.17. The equations (6.15) are relatively easy to be implemented on-chip using simple operational amplifiers as

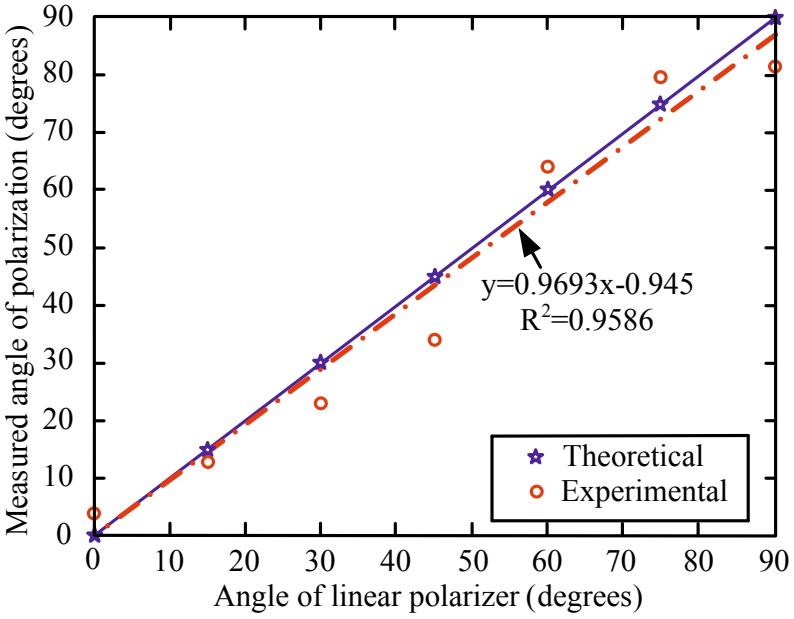


Fig. 6.17 Measured angle of polarization in polarization sense region 2 using equations (6.15)

analog adders or subtractors. The high linearity error is due to the higher metallic wire grid pitch used, limited by the choice of the process technology. With advances in the process technology the pitch can be considerably reduced providing better polarization detection ability and thus a lower linear fit error.

For figures 6.15, 6.16 and 6.17 the transmission axis of the linear polarizer was varied keeping the DC source fixed. The measured azimuthal angles for different angular position of the DC light source in the measurement setup (figure 6.13) for four different measurements are shown in figure 6.18. Similar to the ellipticity angle measurements, the angular position of the light source, and hence the angle of incidence of the light, was varied by 10°, 15° and 30°. A mean correlation coefficient of 0.9904 is obtained between the theoretical and the four measured results [38].

6.7 Real-Time Implementation of Navigation Compass

In section 6.5, it was shown that the polarization Fresnel ratio (PFR) is a good tool to be used as a compass for autonomous agent navigations. To implement the *PFR* calculations in real time using equation (6.11), an analog divider is needed. An analog divider can be implemented either in current mode or in voltage mode using a voltage controlled resistor as shown in figure 6.19 [39]. A current mode analog divider is discussed in this section. To use the current mode analog divider, the voltage at the floating diffusion node of the photodiode needs to be converted to

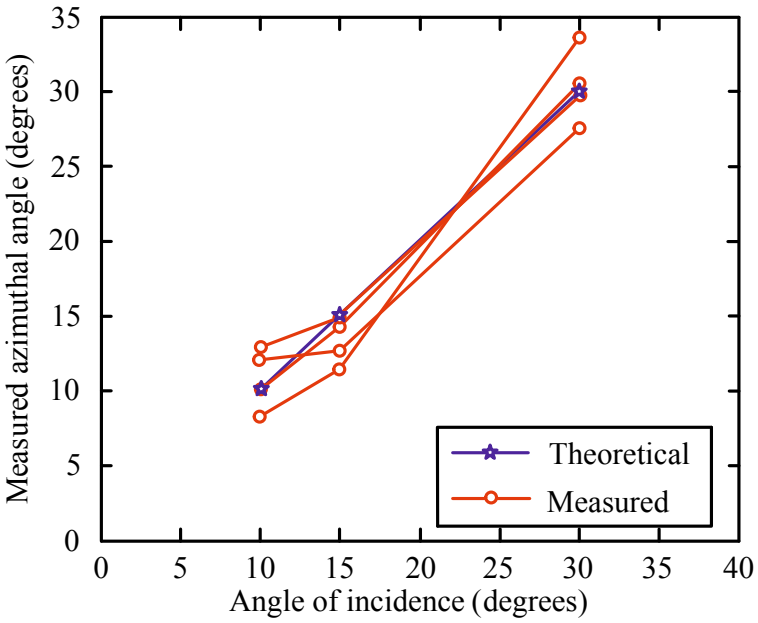


Fig. 6.18 Measurement of incident light ray angle

current. This can be done using a transconductance amplifier or a *V-I* converter. If a conventional *p-n* junction photodiode is used then the photodiode current can directly be copied.

To implement an analog divider, a voltage controlled resistor *(VCR)* is required, an implementation of *VCR* is shown in figure 6.19. If transistors M1 and M2 are biased in triode region then the drain currents I_1 and I_2 are given by:

$$I_1 = I_{in} + I_3 = \frac{K_{n1}}{2}\left(2(V_1 - VTH1)V_{DS1} - V_{DS1}^2\right) \tag{6.16}$$

$$I_2 = \frac{K_{n2}}{2}\left(2(V_2 - VTH2)V_{DS2} - V_{DS2}^2\right) \tag{6.17}$$

where k_{n1} and k_{n2} are the transconductance parameters, V_1, and V_2 are the bias voltages, V_{DS1} and V_{DS2} are the drain to source voltage, V_{TH1} and V_{TH2} are the threshold voltages for the transistors M_1 and M_2 respectively, I_{in} is the additional injected current and I_3 is the drain current of transistor M_3.

The current mirror M_5 and M_6 ensures that I_3 and I_4 have the same value. The current I_4 also passes through the transistor M_2 so that

$$I_3 = I_4 = I_2 \tag{6.18}$$

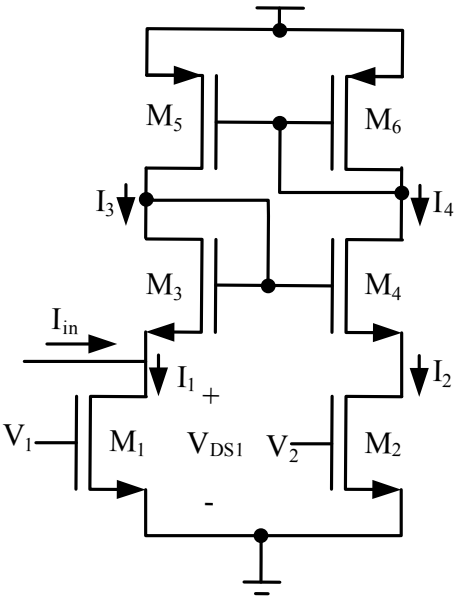


Fig. 6.19 Voltage controlled resistor

Assuming M_3 and M_4 to be perfectly matched and biased in saturation, V_{GS3} is expressed as:

$$V_{GS3} = V_{GS4} = \sqrt{\frac{2I_2}{K_{n4}}} + V_{TH4} \tag{6.19}$$

Also

$$V_{GS4} + V_{DS2} = V_{DS1} + V_{GS3} \tag{6.20}$$

From equations (6.19) and (6.20)

$$V_{DS2} = V_{DS1} \tag{6.21}$$

Substituting equations (6.21), (6.18) and (6.17) in (6.16) and assuming the transistors M_1 and M_2 to be completely matched ($V_{TH1} = V_{TH2} = V_{TH}$ and $K_{n1} = K_{n1} = K_n$), I_{in} can be written as:

$$I_{in} = K_n V_{DS1}(V_1 - V_2) \tag{6.22}$$

From equation (6.22) a voltage controlled resistor between I_{in} and ground can be realized with an equivalent resistance of

$$R_{eq} = \frac{V_{DS1}}{I_{in}} = \frac{1}{K_n(V_1 - V_2)} \tag{6.23}$$

Using the voltage controlled resistor of figure 6.19a current mode divider can be implemented as shown in figure 6.20 [40].

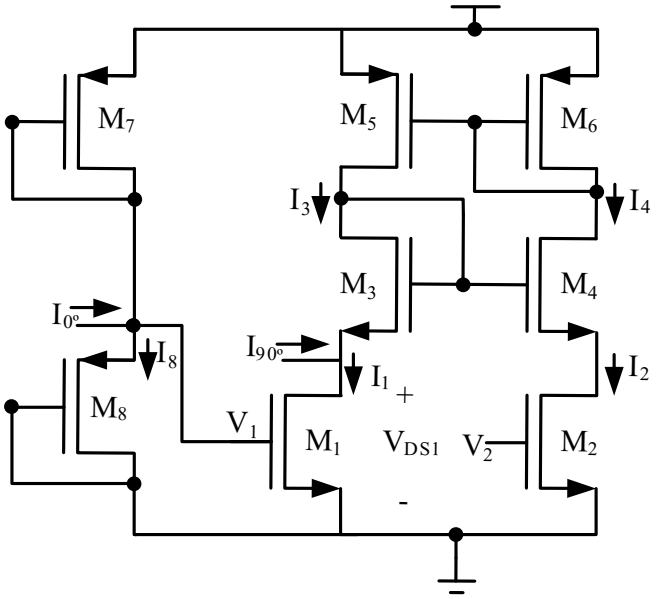


Fig. 6.20 Current mode divider using voltage controlled resistor

Assuming transistors M_7 and M_8 to be perfectly matched ($K_{p7} = K_{p8} = K_p$ and $V_{THP7} = V_{THP8} = V_{THP}$), V_1 can be expressed as:

$$V_1 = \frac{I_{0^\circ}}{2K_p(-V_{THP})} \tag{6.24}$$

Substituting equation (6.23) in (6.21) and grounding V_2

$$V_{DS1} = \frac{2K_p(-V_{THP})}{K_n} \times \frac{I_{90^\circ}}{I_{0^\circ}} \tag{6.25}$$

According to equation (6.25) the output V_{DS1} is proportional to the ratio of the two input currents I_{90° and I_{0° which are the outputs of the two pixels sensitive to 0° and 90° sensitive pixels in the designed image sensor. This allows calculation of the *PFR* in continuous time.

Besides the requirement of an analog divider, the simplified Stokes parameter calculations in equation (6.15) also need a subtraction of the two pixel values which can be easily obtained using the circuit shown in figure 6.21.

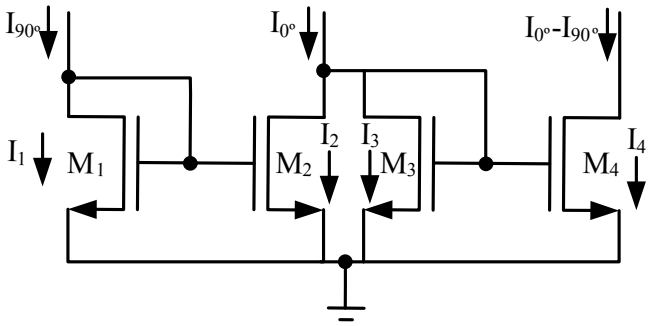


Fig. 6.21 Current mode subtraction circuit

If transistors M_1, M_2, M_3 and M_4 are in saturation, and assuming they are perfectly matched the currents in these two transistors are given by

$$I_1 = I_{90^\circ} = I_2 = \frac{K_n}{2}(V_{GS1} - V_{TH})^2 \tag{6.26}$$

$$I_3 = I_{0^\circ} - I_{90^\circ} \tag{6.27}$$

Since M_3 and M_4 acts as a current mirror the current I_3 is copied by the transistor M_4 and the output current is the difference of the two input currents from the two pixels sensing 0° and 90° polarization.

6.8 Summary

- There are two types of spatial representation created by navigation, one is cognitive map or survey representation and the other is route representation or spatial representation.
- The navigation can be broadly classified as geocentric and egocentric based on the cognitive mapping.
- Human use geocentric navigation wherein the visual clues such as landmarks play an important role. Insects use egocentric navigation, wherein the direction of travel is determined from the position of the sun or from skylight polarization.
- The egocentric navigation is best suited for autonomous agents like robotic navigation as it does not involve large frame to frame image computations to determine the distance and direction of travel.
- Path integration is the basis of vector navigation in egocentric forms of navigation, it requires knowledge of the direction of travel and the distance traveled.

- To determine the direction of travel, a reference direction is needed. Insects use the position of the sun as the reference and compute their travel direction with respect to the position of the sun. The movement of the sun is constant and is approximately 1°every 4 minutes.
- The skylight polarization is also used to determine the reference direction. The skylight polarization is dependent on the position of the sun.
- The skylight polarization pattern varies in a systematic fashion both in plane (*e*-vector) and degree of polarization, according to the position of the sun.
- There exists an indirect relationship between the degree of polarization and the skylight intensity: for sun positions near the zenith, the degree of polarization is quite high while at the horizon the degree of polarization is weaker.
- The principle of using polarization antagonism prevalent in the compound eye of insects as a polarization based compass for autonomous agents has been successfully modeled by various researchers.
- From the micro-polarizer orientations, intensities can be used to calculate the degree of polarization and polarization Fresnel ratio for varying angle of the incoming light which can be used as a compass. The positional information can also be retrieved using Stokes parameters.
- The sun position detection using polarized light would have an advantage over conventional sun position detection sensors as it would work even when the sun is not visible.
- The Stokes parameters are used to compute the ellipticity and azimuthal angles of the perceived light ray, which is correlated to the incoming light ray direction.
- The simplified computational algorithms allow on-chip processing which would result in miniaturized navigational sensors. Such navigational sensor would be independent of the visual cues and use natural light to determine the directional reference.

References

[1] O'keefe, J., Nadel, L.: The Hippocampus as a Cognitive Map. Oxford University Press (1978) ISBN: 0198572069

[2] Thorndyke, P., Hayes-Roth, B.: Differences in spatial knowledge acquired from maps and navigation. Cognitive Psychology 14, 560–589 (1982)

[3] Hafner, V.: Adaptive Navigation Strategies in Biorobotics: Visual Homing and Cognitive Mapping in Animals and Machines. Shaker Verlag (2004) ISBN: 3832228578

[4] Wehner, R., Michel, B., Antonsen, P.: Visual navigation in insects: coupling of egocentric and geocentric information. Journal of Experimental Biology 199, 129–140 (1996)

[5] Moravec, H.: Obstacle avoidance and navigation in the real world by a seeing robot rover. Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotic Institute (1980)

[6] DeSouza, G., Kak, A.: Vision for mobile robot navigation: A survey. IEEE Transactions on Pattern and Machine Intelligence 24, 237–267 (2002)

[7] Lambrinos, D., Möller, R., Labhart, T., Pfeifer, R., Wehner, R.: A mobile robot employing insect strategies for navigation. Robotics and Autonomous Systems 30(2), 39–64 (2000)
[8] Lambrinos, D., Maris, M., Kobayashi, H., Labhart, T., Pfeifer, R., Wehner, R.: An autonomous agent navigating with a polarized light compass. Adaptive Behaviour 6(1-2), 131–161 (1997)
[9] Usher, K., Ridley, P., Corke, P.: A camera as a polarized light compass: preliminary experiments. In: Proceedings of Australian Conference on Robotics and Automation, pp. 116–120 (2001)
[10] Müller, M., Wehner, R.: Path integration in desert ants, Cataglyphis Fortis. Proceedings of National Academy of Sciences 85, 5287–5290 (1988)
[11] Chen, Z., Birchfield, S.: Qualitative vision based mobile robot navigation. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 2686–2692 (2006)
[12] Moini, A.: Vision chips. Kluwer academic publishers, Boston (2002) ISBN: 0792386647
[13] Müller, M., Wehner, R.: The hidden spiral: systematic search and path integration in desert ants, Cataglyphis Fortis. Journal of Comparitive Physiology A-Sensory Neural and Behavioural Physiology 175, 525–530 (1994)
[14] Wehner, R., Müller, M.: The significance of direct sunlight and polarized skylight in the ant's celestial system of navigation. Proceedings of the National Academy of Sciences, USA 103(33), 12575–12579 (2006)
[15] Santschi, F.: Review Suisse Zoology, vol. 19, pp. 303–338 (1911)
[16] Liebe, C.: Solar compass chip. IEEE Sensors Journal 4, 779–786 (2004)
[17] Chang, Y., Kang, S., Lee, B.: High accuracy image centroding algorithm for CMOS based digital sun sensors. In: Proceedings of IEEE Sensors Conference, pp. 329–336 (2007)
[18] Xie, N., Theuwissen, A., Wang, X.: A CMOS image sensor with row and column profiling means. In: Proceedings of IEEE Sensors Conference, pp. 1356–1359 (2008)
[19] Bjorn, L.: Photobiology: the science of life and light, p. 18. Springer publications (2007) ISBN: 9780387726540
[20] Coulson, K.: Polarization and Intensity of Light in the atmosphere, Deepak, Hampton, VA, p. 2 (1988) ISBN: 0937194123
[21] Smith, G.: The polarization of skylight: An example from nature. American Journal of Physics 75(1), 25–35 (2007)
[22] Richardson, R., Hulburt, E.: Sky-brightness measurements near bocaiuva, brazil. Journal of Geophysics Research 54, 215–227 (1949)
[23] Wehner, R.: Desert ant navigation: how miniature brains solve complex task. Journal of Comparitive Physiology A-Sensory Neural and Behavioural Physiology 189, 579–588 (2003)
[24] Seidl, T.: Ant navigation and path finding. In: Proceedings of the 2nd ACT Workshop on Innovative Concepts, vol. 31(4), pp. 102–110 (2008)
[25] Pieron, H.: Du role du sens musculaire dans l'orientation de quelques espéces de fourmis. Bulletin of Institute of General Psychology 4, 168–186 (1904)
[26] Cornetz, V.: Trajets de fourmis et retours au nid. Mémoire De l'Institute Géneral, Psychologie 2, 1–67 (1910)
[27] Mittelstaedt, M.L., Mittelstaedt, H.: Homing by path integration in a mammal. Naturwissenschaften 67, 566–567 (1980)
[28] Wehner, R.: The ant's celestial compass system: spectral and polarization channels. In: Orientation and Communication in Arthropods, pp. 145–185. Birkhäuser Verlag, Basel (1997)

[29] Labhart, T., Meyer, E.: Detectors for polarized skylight in insects: a survey of ommatidial specializations in the dorsal rim area of the compound eye. Microscopy Research and Technique 47, 368–379 (1999)
[30] Herrling, P.: Regional distribution of three ultrastructural retinula types in the retina of cataglyphis bicolor fabr (formicidae, hymenoptera). Cell Tissue Research 69, 247–266 (1976)
[31] Labhart, T.: How polarization-sensitive interneurons of crickets see the polarization pattern of the sky: a field study with an opto-electronic model neuron. Journal of Experimental Biology 202, 757–770 (1999)
[32] Bernard, G., Wehner, R.: Functional similarities between polarization vision and color vision. Vision Research 17, 1019–1028 (1977)
[33] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.: Integrated polarization analyzing CMOS image sensor for autonomous agent navigation using polarized light. In: Proceedings of IEEE International conference on Intelligent Systems, pp. 224–229 (2010)
[34] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.: Integrated polarization analyzing CMOS image sensor for detecting incoming light ray direction. In: Proceedings of IEEE Sensors Application Symposium, pp. 194–199 (2010)
[35] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.: Biologically inspired autonomous agent navigation using stokes parameters and an integrated polarization analyzing CMOS image sensor. In: Proceedings of Eurosensors XXIV Conference, vol. 5, pp. 673–676 (2010)
[36] Gruev, V., van der Spiegel, J., Engheta, N.: Integrated polarization image sensor for cell detection. In: International Image Sensor Workshop (2009)
[37] Gruev, V., Perkins, R.: A 1 mpixel ccd image sensor with aluminum nanowire polarization filter. In: Proceedings of IEEE International Symposium on Circuits and Systems, pp. 629–632 (2010)
[38] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.J.P.: Integrated polarization analyzing CMOS for navigation and incoming light ray direction. In: IEEE Transactions of Instrumentation and Measurement (accepted for publication)
[39] Liu, W., Liu, S.: CMOS tunable 1/x circuit and its applications. Transactions on Fundamentals of Electronics, Communications and Computer Sciences E-86A, 1896–1899 (2003)
[40] Liu, W., Liu, S., Wei, S.K.: CMOS current-mode divider and its application. IEEE Transactions on Circuits and Systems-II 52(3), 145–148 (2005)

Chapter 7
Motion Detection and Digital Polarization

Flying insects have extraordinary visual capabilities. Their ability to detect fast motion in the visual scene and avoid collision using low level image processing and little computational power makes their visual processing interesting for real time motion/collision detection in machine vision applications.

In this chapter, some bio-inspired models of motion and collision detection based on differential imaging and correlation will be presented. Section 7.1 describes the process of obtaining motion from a sequence of images. The conventional models for motion detection using differential imaging, background subtraction and optical flow are also briefly introduced. Motion detection using differential imaging is described in section 7.2. Section 7.3 describes the most popular bio-inspired elementary motion detector. It also discusses on the possible simplification of elementary motion detection using one-dimensional binary optical flow. The spatial summation of differential one-dimensional binary optical flow is shown to be able to detect both horizontal and vertical object motion. The partial charge transfer method to increase the dynamic range and the background illumination invariance are described in section 7.4. A digital representation of polarization is presented in section 7.5, where the one-dimensional binary optical flow is shown to vary with the polarization angle of the incoming light rays.

7.1 Motion Detection

Motion is usually determined from image sequences. The spatiotemporal image sequences can be represented using the plenoptic function. The plenoptic function was introduced by Adelson and Bergen [1] to describe all the information available to an observer at any point in space and time, when a light ray passes through the imaging device. In its most general form, the plenoptic function is a seven-dimensional function given by the equation (7.1)

$$P = P(\theta, \varphi, t, \lambda, V_x, V_y, V_z) \tag{7.1}$$

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 215–245.
DOI: 10.1007/ 978-3-642-34901-0_7 © Springer-Verlag Berlin Heidelberg 2013

where (V_x, V_y, V_z) is the position of the imaging sensor in three-dimensional space, λ is the wavelength of light, (θ, φ) represents the azimuth and elevation angles that index the viewable rays and t is time. In a pinhole camera only one sample of the viewpoint is available at any given instance, thus the plenoptic function can be adapted to standard Cartesian parametrization of the rays *(x,y)*, where *x* and *y* are the spatial coordinates in the image plane.

$$P = P(x, y, t, \lambda) \tag{7.2}$$

Though multispectral imaging has started to become popular, most motion analysis are done for a single wavelength and thus the plenoptic function can be rewritten as

$$P = P(x, y, t) \tag{7.3}$$

The plenotic function of a time sampled set of images or "snapshots" is denoted by

$$P = P(x, y, \{t = i, i + \delta t, ..., i + n\delta t\}) \tag{7.4}$$

where *i* is the time instance of an image capture, δt is the time interval between two consecutive images and *n* is total number of images captured. In machine vision, the variations in the plenotic function across sequences of images are used to calculate the image velocity or motion. One of the measurable parameters of a point object which is assumed not to change with time for changing plenoptic function is the intensity or the brightness of the object given by *I(x,y,t)*, where *(x,y)* represents the spatial x and y dimensions and t denotes time [2]. It is thus useful to combine the geometrical description of the scene with brightness information for motion detection [3]. The brightness constancy equation assuming the intensity of a point to remain constant as it moves δx, δy for a time interval δt can be written as

$$I(x - \delta x, y - \delta y, t - \delta t) = I(x, y, t) \tag{7.5}$$

Taylor series expansion of equation (7.5) can be written as:

$$I(x - \delta x, y - \delta y, t - \delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t + \text{higher order terms} \tag{7.6}$$

Neglecting the higher order terms in equation (7.6) and using the brightness constancy equation [1] we get

$$\frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t = 0 \tag{7.7}$$

where $\partial I/\partial x$, $\partial I/\partial y$ and $\partial I/\partial t$ are the derivatives of the image *(x,y,t)* in the corresponding directions and thus can be represented by I_x, I_y and I_t. Equation (7.7) can be rewritten as:

$$I_x \delta x + I_y \delta y + I_t \delta t = 0 \tag{7.8}$$

Dividing all the terms in equation (7.8) by δt we get:

$$I_x u + I_y v + I_t = 0 \tag{7.9}$$

where *u(x,y,t)* and *v(x,y,t)* are the horizontal and vertical components of the motion respectively.

Equation (7.9) relates the image velocity to the spatiotemporal derivative of the image at a particular location and is commonly referred to as the motion constraint equation.

Based on the variation in the intensity obtained from the projection of the plenoptic function on the image sensor from the moving object, three conventional approaches are used to detect motion: temporal differencing [4]; background subtraction [5], [6]; and optical flow [2].

Temporal differencing is based on frame difference, and attempts to detect moving regions by making use of the difference of consecutive frames (two or three) in a video sequence. This method is highly adaptive to dynamic environments, but generally does a poor job of extracting the complete shapes of certain types of moving objects [7].

Background subtraction uses a model of the background and compares the current image with the reference image to separate the background and foreground [8]. The main disadvantage of background subtraction method is that both the background scene and the camera are required to be stationary when this method is applied. They are also extremely sensitive to dynamic scene changes due to background illumination changes.

When either the optical scene or the camera is in motion, optical flow is usually used. An object in motion may exhibit both translation and rotational velocities, which are usually projected as the movement of brightness patterns on the image plane. As per the smoothness constraints, the corresponding points in two successive frames should not move more than a few pixels. Thus in an uncertain environment the camera motion or background changes should be relatively small and thus motion can still be detected with optical flow. Optical flow estimation methods can be classified into three main groups: differential methods; matching-based methods and frequency/phase-based. All the three methods consist of three basic components: pre-filtering, local motion estimation and integration over the field of view. Pre-filtering or smoothing of the image data with a low-pass or band-pass filter allows extracting the signal of interest thus enhancing signal-to-noise ratio.

The local motion estimation is done using spatiotemporal derivatives to measure the velocity component and the integration over the field of view produces a two-dimensional flow field of the moving object.

The differential methods include the gradient based models that determine two-dimensional velocities of the moving object employing first or second order spatiotemporal derivatives of the image sequence. The velocity is obtained by dividing the temporal derivative of local luminance by its spatial derivative. The major problem with gradient based methods is that if the spatial derivative is small, the noise in the temporal derivative is amplified and the velocity is poorly defined. Thus due to the differentiation of the image sequence and their susceptibility to errors under noisy conditions they are usually not preferred in low signal-to-noise ratio regimes [9].

Matching-based methods include feature-based and region-based methods. They are also known as block-based, area-based and correlation based methods. These methods usually use techniques to either maximize cross-correlation or minimize differential error. Feature based methods locate and trace identifiable features of the image over time while region based methods try to locate a delineated region in consecutive frame within a search space. To match the feature or the regions in subsequent images either probabilistic approaches like Kalman filters and Monte Carlo localizations or neural networks are used. Matching based methods are robust to large motion and brightness variations however these methods are found to be accurate only at high velocities and it is difficult to estimate sub-pixel displacements [10]. Further correlation based models are normally very sensitive to the amount of data involved because these algorithms are mainly based on byte-level operations of the whole image. These methods involve an operation of all the pixels in the image even though most pixels may not have changed from one frame to another, severely limiting the bandwidth and speed or motion detection for real-time applications.

Frequency/phase based methods use local energy or phase information to determine the velocity of the moving object. These techniques for determining image motion rely on the phase behaviour of arrays of band-pass filters or changes in the output energy of the velocity tuned filters. These filters decompose the input signal according to scale, speed and orientation. However these methods are still susceptible to noise and discontinuity limitations like the gradient based methods. Additionally the outputs of these methods are limited by the design of the filters used [10].

7.1.1 Motion Detection - Models

The problem of real-time motion detection and tracking is an important issue in artificial vision. The main constraints for real-time implementation of these algorithms are the large amount of data to be processed and the high-computational cost of the algorithms employed. To solve the problems, analog VLSI chips employing early vision processing of the optical scene are becoming popular [11], [12], [13], [14]. They employ simple, low accuracy operations at each pixel in an image or sequence

of images, resulting in a low-level description of a scene useful for higher level machine vision applications. These often results in compact, high speed and low power solutions.

The motion detection can be done either on a pixel level or on a region level. The pixel level analysis shows whether the pixel is stationary or transient by observing its intensity value over time. Region analysis deals with the agglomeration of groups of pixels into moving and stationary regions.

It is usually known that the pixel's intensity value displays three characteristic profiles as shown in the figure 7.1. When an object moves through the pixel it displays a profile that exhibits a step change in intensity, followed by a period of instability, then another step back to the original background intensity. When the object moving through the pixel stops at the pixel location, it exhibits a change in intensity, followed by a period of instability and then settling to the new intensity levels. With variations in the ambient lighting, the intensity changes smoothly with no large steps [15]. To know the state of the object, the nature of the pixel intensity profile is important. To interpret the meaning of a step change (for example object passing through, stopping at, or leaving the pixel), one needs to observe the intensity curve re-stabilizing after the step change. This introduces a time-delay into the process. For detection of fast motion this time delay has to be minimized.

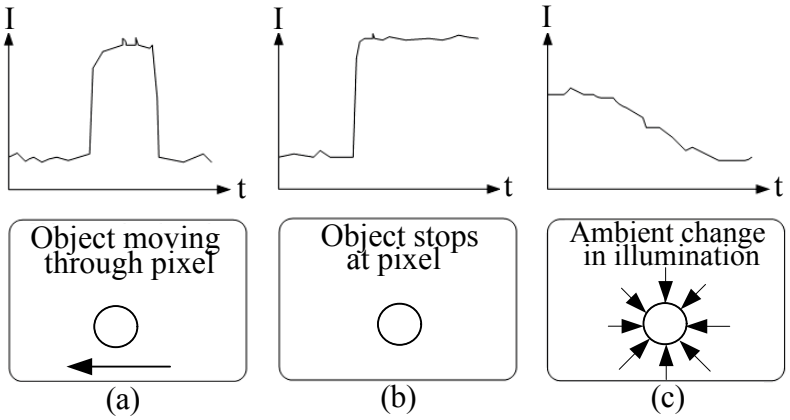


Fig. 7.1 Characteristic pixel intensity profiles for (a) object moving though the pixel, (b) object stopping at the pixel and (c) variations in the ambient illumination changes

In motion detection algorithms where the image motion is analyzed by sending the complete frame out of the image sensor, a very high data rate is required to keep the time delay small. The common method employed to obtain differential images is to use the CCD or CMOS image sensor in high speed frame mode combined with fast readout, a frame buffer to store the frame and a digital subtractor to generate the differential image. The high frame rate is needed to prevent temporal aliasing, which prevents the output signal from being a smooth function of the image parameters thus affecting the efficiency of motion detection algorithms. Furthermore a long time between two image captures will give a difficult correspondence problem in

a dynamic scene. Such an image sensor is highly complex and consumes lot of power. Thus such motion sensing algorithms though in use, for example in MPEG encoders, are not very convenient for real-time motion applications in machine vision. These algorithms for fast motion detection become design overkill for slow motion detections where the time delay can be large.

Temporal differencing at a pixel level rather than at a frame level can help to decrease the data coming from the camera as only the data from the pixel with changed states will be sent out [12], [16]. The focal plane computations are also free from temporal aliasing, which is usually a problem in motion detection algorithms. Further temporal differencing is useful for ambient light suppression. Two methods are discussed, one analog and one digital, in the subsequent sections that can be employed for speeding up low level motion detection without much complexity and power consumption. These algorithms are based on pixel changes instead of full image processing and thus improve performance. The pixel level processing provides an extreme data rate compression; it has a high data rate in (image sequence) and low data rate out (motion).

7.2 Motion Detection - Differential Imaging

In differential imaging two consecutive images are stored in the pixel as an analog voltage and a difference signal is obtained during the readout. Differential imaging can be used for machine vision applications like motion detection, object tracking or object recognition etc. All the stationary objects with constant illumination will be ignored by the imaging system.

An active differential sensing method is described here, with focal plane computations which need minimal external active components. The subtraction of the two images captured at varying exposure time is done pixel by pixel. The process is repeated for each image line and thus a differential image is obtained without a requirement for an off-chip computation.

The designed image sensor explained in chapter 3 can operate in two modes: double differential mode *(DDS)* and differential imaging mode *(DI)* using the two available analog memories in the pixel. The operation of the image sensor in the *DDS* mode was shown in section 3.6.5.1. The pixel diagram with the two analog memories and the operational timing diagram of the *DI* mode of the pixel are shown in the figure 7.2.

In the *DI* mode the sensor takes two samples and subtracts one from the other. The first sample is taken after an integration time T_1 while the second sample is taken after an integration time T_{FT} and are stored in the two analog memories in the pixel. The imaging array operates in global shutter mode as the signals after the integration time is transferred to the memories at the same time for all the pixels of the imaging array. However while reading out, each row is read at a time. The two signals stored in the pixel are subtracted at the column level of the analog signal chain shown in figure 3.33. The subtraction of the two images reduces the *FPN* and since there is no reset between the two image captures the reset noise is the same in both images and it is removed with subtraction.

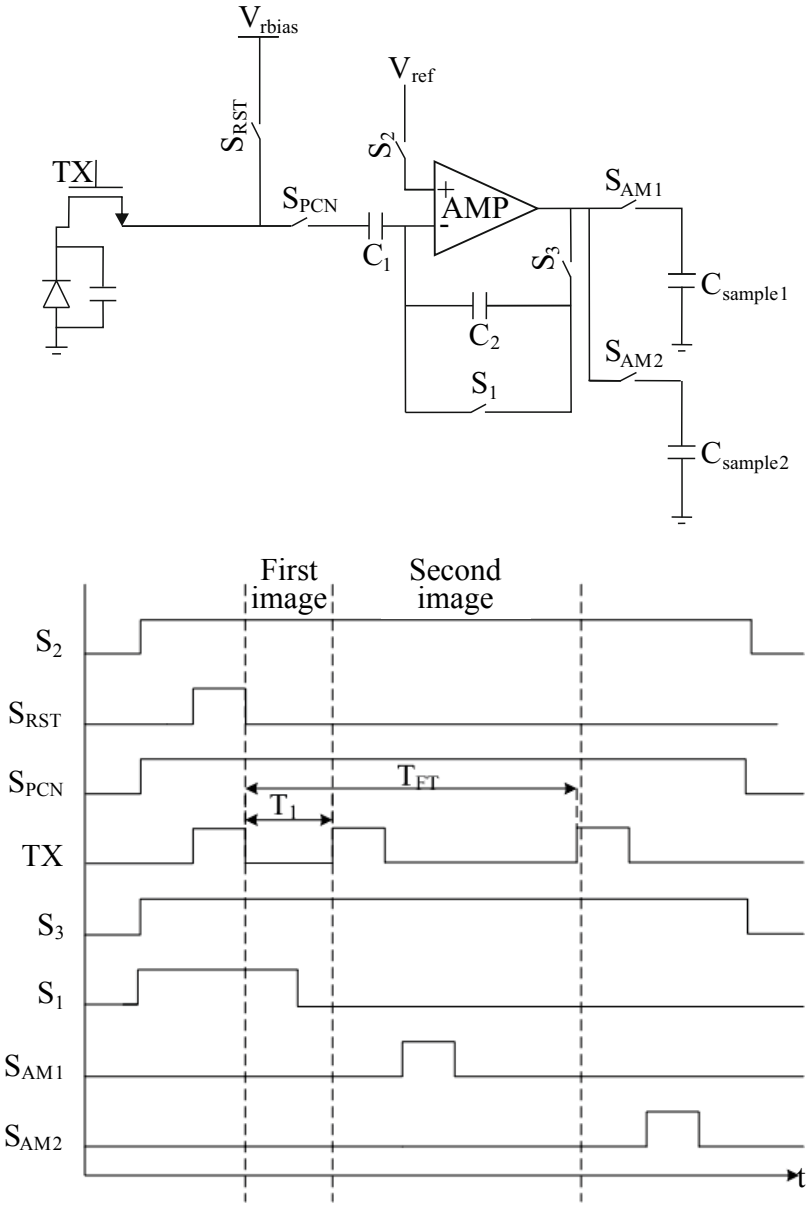


Fig. 7.2 Differential imaging mode

Figure 7.3 shows an example of motion detection, when the object moves to a pixel position and stops, similar to the situation shown in figure 7.1(b).

When the exposure time of the two samples is the same $(T_1 = T_{FT}) = 20ms$, a black image is observed when there is no change in the background as shown in

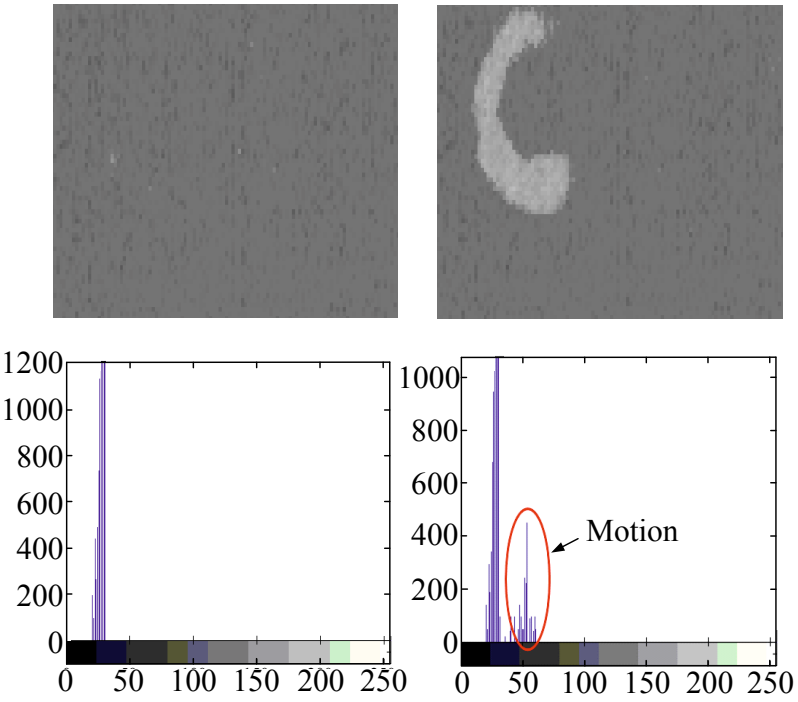


Fig. 7.3 Motion detection using differential mode of operation

figure 7.3(a). When a DC light source is switched on (to resemble motion), there is a change in the background and the change is visible in the differential image 7.3(b). The histograms shows the gray values of the pixels. When there is no motion, most of the pixels have lower gray values while when there is motion pixels start to get higher gray values.

For motion detection all the static objects or DC light sources in the background have to be removed and only the AC change is to be detected. Thus a very high Common-Mode Rejection Ratio *(CMRR)* is desired to remove the common mode or DC signal in subsequent images of a scene. Common-mode rejection ratio of a device is the measure of the tendency of the device to reject common input signal. A high *CMRR* is desired for applications where the signal of interest is a small voltage fluctuation superimposed on a large voltage offset or when the required information is contained in the voltage difference of two signals.

The definition of the *CMRR* of the differential image is simplified from the conventional definition of *CMRR* and defined to be the change in the common mode signal divided by the change in the differential signal due to the change in the common mode signal [17]. A high *CMRR* is desired to have the same effect when two white images or two black images are subtracted from each other. The maximum voltage swing obtained after *DDS* for the designed sensor is 488mV as seen in section 3.7.2.1, for dark to saturation exposure of the sensor. The output

voltage swing obtained when operating in the differential mode ($T_1 = T_{FT}$) is 2.44mV for varying illumination from dark till the sensor saturation. The *CMRR* thus obtained is 200 or 46dB which is higher than the *CMRR* of 100 reported in [17].

The changes in the differential signal when the normal exposure period (T_{FT}) is held constant at 20ms while the short integration time (T_1) is varied is shown in the figure 7.4.

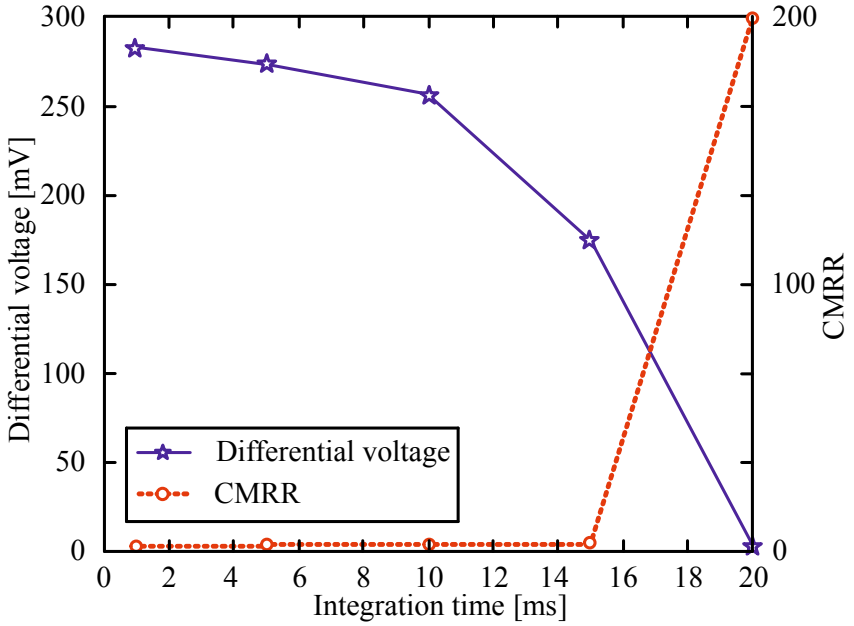


Fig. 7.4 Variation in the differential swing and CMRR with changes in the exposure time

A non-linear behavior is seen when T_1 is very low compared to T_{FT}. If the ratio of the two integration periods is defined as $R_{DE} = T_{FT}/T_1$, then it is observed that for larger values of R_{DE} the non-linearity is quite high. The accumulated electrons for a given illumination condition decrease with the increase in R_{DE} due to shorter accumulation period (The accumulated electrons are proportional to $1/R_{DE}$). When the accumulated electrons are low, the error in transferring these accumulated electrons from the photodiode to the floating diffusion node is higher. Thus the non linearity increases with increase in R_{DE}. The results are very much in agreement with the theoretical studies of [18]. The *CMRR* practically remains low for all values of T_1. When T_1 is equal to T_{FT}, it reaches the maximum value of 200.

7.3 Motion Detection - Optical Flow

A one-dimensional motion can either be in a horizontal direction or in a vertical direction. Vertical motion of an object towards the sensor leads to a collision

and thus needs to be prevented. The visual guided collision avoidance has been extensively studied using conventional cameras and digital processing devices. However conventional cameras are not suitable for real time applications as they rely on motion estimation using sequential images.

It is well known that flying insects are able to detect obstacles in their flying path efficiently with little computational power by using optical flow. Optical flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and the scene. Optical flow usually contains information about self motion and distance to potential obstacles and thus it is very useful for navigation [19]. In section 7.1, the three methods of obtaining optical flow for motion detection have been summarized. These methods are however broadly equivalent. In biological motion vision, the correlation based model or the gradient detector model are often used to account for the direction selectivity [20]. The correlation based model is shown to have significant advantages over the gradient based model in regimes where signal-to-noise ratio is low and detector noise is of concern.

The circuit of each pixel of the designed image sensor contains a comparator to detect the difference between the integrated charge from the photodiode and an external threshold voltage (chapter 3). This allows the generation of binary optical flow similar to the effect of *"flickering"* in the eyes of insects. The binary optic flow is the change in the digital pixel values in response to objects motion in the focal plane of the image sensor. The generated optical flow can be used to detect motion both in vertical (section 7.3.1) and horizontal (section 7.3.2) direction with minimal processing and hardware.

7.3.1 Motion in Vertical Direction - Collision Detection

In this section vertical motion in the focal plane is considered to estimate the time to collision of a moving object with the image sensor using binary optic flow. Reliable estimation of the time to collision between two moving objects is very important in many applications such as autonomous agent navigation or predictive crash sensors for autonomous safety systems. Currently existing non-biologically inspired collision avoidance systems use a CCD/CMOS camera and digital processing devices to detect the approaching object. Such a collision detection system is not suitable for compact real-time computations as it requires large amount of computations.

7.3.1.1 Elementary Motion Detector (EMD)

The most popular bio-inspired visual guided collision avoidance approach uses the correlation-type elementary motion detector *(EMD)*, first proposed by Hassenstein and Reichardt to compute the optical flow [21]. They have been used to explain directional selective motion in a wide variety of insects, birds and mammals including humans [10]. This model is very well established and often used in

bio-inspired robots [22], [23] and [24]. The *EMD* correlates the response of one photoreceptor to the delayed response (inhibitory response) of an adjacent photoreceptor, both looking in the same direction as shown schematically in figure 7.5.

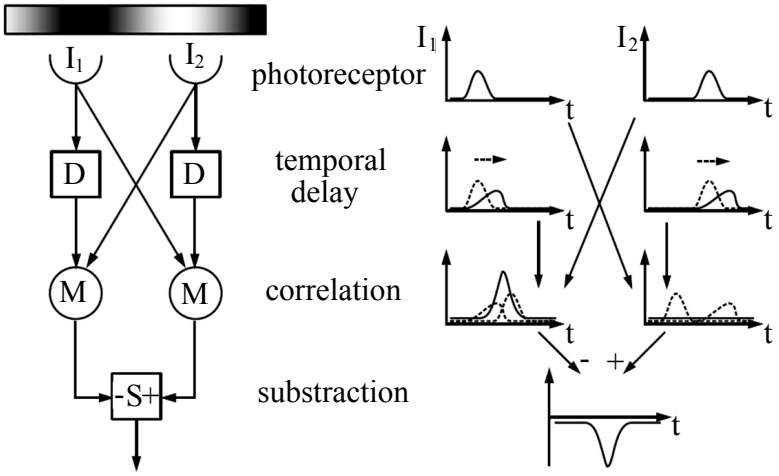


Fig. 7.5 The Hassenstein-Reichardt or correlational elementary motion detector (EMD) [25]

The elementary motion detector consists of two spatially separate inputs to measure the changes across space, temporal filters (delay) to measure the changes across time and a comparator to evaluate spatial and temporal changes. The coincidence of the original signal from one point in space and the delayed signal from the neighboring point in space leads to a positive output signal. The final output signal of the *EMD* is a transient response obtained by subtracting two images. The output image is sensitive to temporal changes in intensity of light, enhancing the directional properties and motion detection by rejecting the effects of temporal contrast not resulting from motion. The motion detection by subtracting two images is better suited than the motion detection using derivative which amplify the noise.

The two separate inputs of the *EMD* model achieve what is referred to as "temporal decorrelation" or "transient enhancement" [23], which is a way of removing the redundant information from the photoreceptor signal before further processing to determine motion. The temporal decorrelation is a filtering mechanism, in which the DC light (mean ambient light) which doesn't have any motion information will be subtracted and AC light (which corresponds to motion) would be allowed to the next stage for further processing.

Existing implementations of the *EMD* use a complex circuit with many active and passive components in order to obtain the inhibition of the signal and the correlation. Further because the basic collision detectors try to measure the velocity of the approaching object, they need dedicated processing blocks for velocity calculation from optical flow field. Additionally, the outputs of the *EMD* are not invariant to the changes in the background illumination and their responses are not only

proportional to velocity changes but are strongly affected by the contrast and the spatial frequency components of the scene [26]. Because the *EMD* multiples its two input, it has quadratic dependence on the contrast and is thus not suitable for low contrast motions. This problem of quadratic dependence can be solved by introducing static expansive non-linearity in the channels [23] however this makes the model more complex. Further because the time delay in the two channels is constant, the *EMD* may infer a wrong conclusion if the time intervals of the moving object between the two points is not the same as the delay.

The standard output of a collision detector using an *EMD* is a response which stays low when the object is far enough and peaks before collision for an approaching object and then collapses to low values again [23]. The peaking of the collision detection algorithm depends on the delay in the delaying channel and the time an object takes to appear in the neighboring channel as well as the variations in the background illuminations [27]. Thresholding of this response to ascertain collision is thus difficult.

7.3.1.2 Proposed EMD Model for Collision Detection

The correlation based models are usually computationally intensive and have difficulties in selecting reliable estimates. In order to have compact and low-power real time autonomous motion detection systems, the estimation of the optical flow has to be simplified. This means that rough qualitative properties of the optical flow are more desirable for efficient collision detection than accurate target distance estimation [19]. The collision avoidance maneuvers in insects can be explained in terms of perception of looming stimuli or expanding images. The landing behavior of insects and the saccade (rapid turns) exhibited by example flies are believed to be triggered by image expansion as detected by an array of local motion detectors [28]. Collision detection using expanding images is also relatively independent of the spatial structure of the object being approached.

To understand the plenoptic function of expansion of images the motion constraint equation (7.9) is used. Since in collision detection the motion is assumed to be only in the vertical direction of the focal plane, the motion constraint equation can be rewritten as

$$I_y v + I_t = 0 \tag{7.10}$$

Equation (7.10) shows that the time of impact or time of collision of the object is directly proportional to the reciprocal of velocity of the object. Equation (7.10) can be better understood using figure 7.6, which shows an image of an approaching object of diameter *D* at a constant velocity *V* along the optical axis. The distance between the lens and the object is *d(t)* while the focal length of the lens is *f*.

The diameter of the obtained image and its derivative with respect to time is given as:

$$a(t) = \frac{f \times D}{d(t)} \tag{7.11}$$

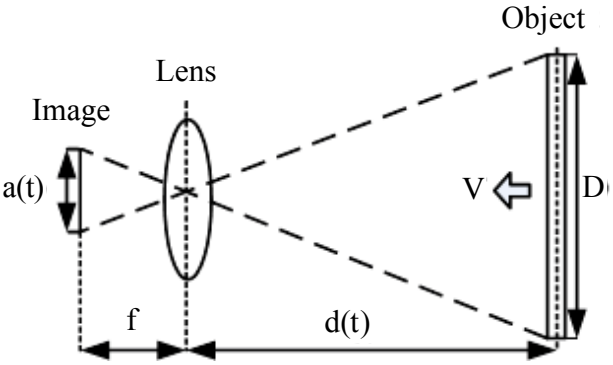


Fig. 7.6 Perception of approaching objects

$$\frac{d(a(t))}{dt} = -\frac{f \times D}{d(t)^2} \tag{7.12}$$

The size of the image *a(t)* in equation (7.11) increases with a decreasing distance *d(t)* and vice versa. The change in the image size affects the optical flow perceived by the image sensor.

To simplify the optical flow generation algorithm, the binary output of the pixel is used. The binarization of the data in the pixel is explained in section 2.6.2.3. In the collision detection experiments, the image sensor is held stationary so that the optical flow is always generated by the motion of the object in the visual field. Figure 7.7 shows the variation in the light spot (approaching object). As the object moves closer to the image sensor, the image size (spot size on the imaging plane) grows or the optical flow expands. With the expanding of the optic flow the intensity profile of the pixel will also increase (figure 7.1(c)), and more pixels will have an output voltage higher than the reference voltage, and a digital '1' will be stored in the *SRAM* cells.

Fig. 7.7 Perception of approaching objects

From the digital images, the percentage of active high pixels for a given illumination condition can be computed as the ratio of total active high pixel to the total number of pixels in the array:

$$\%\ \text{of active high pixels} = \frac{\text{Total active high pixels}}{\text{Total number of pixels}} \tag{7.13}$$

Equation (7.13) represents the one-dimensional binary optical flow. The percentage of the active high pixels will increase with the approaching bright object as predicted by equation (7.11). The measured variations in the percentage of active high pixels with the variation in the distance of the light source for single image capture is shown in figure 7.8. It shows that when the light source approaches the image sensor, the optical flow, which is the variation in the intensity with motion, causes more pixels to become active thus increasing the percentage of active high pixels.

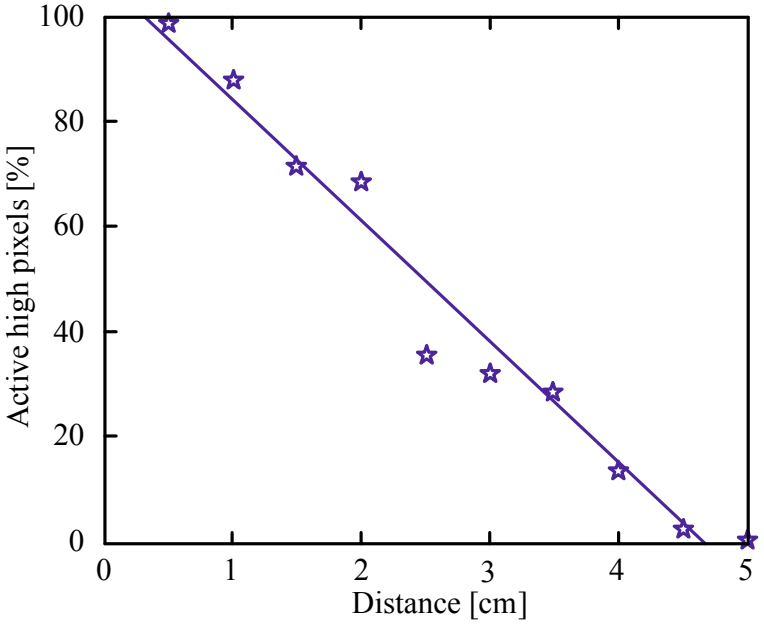


Fig. 7.8 One-dimensional binary optical flow variation with approaching object

One of the major requirements of motion detection using correlation models is temporal decorrelation. Temporal decorrelation refers to the process of reducing the autocorrelation with a signal in time domain, this helps in removing redundant information in images separated in time. Temporal decorrelation can be obtained using differential imaging, where two samples are spaced in time. The differential image is generated using partial charge transfer, where the integrated charge at the photodiode capacitance is transferred to the *FD* node multiple times in one frame (chapter 3). Figure 7.9 shows the algorithm used.

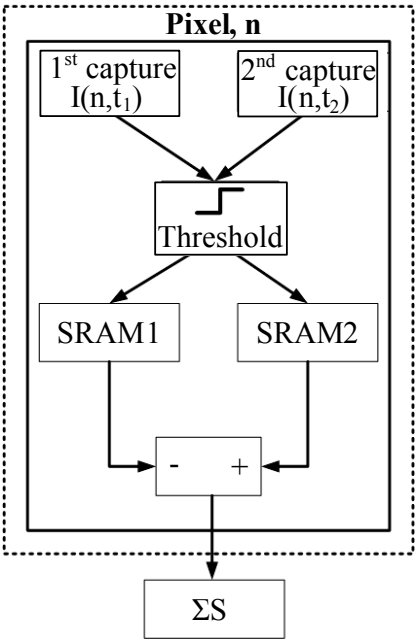


Fig. 7.9 Modified EMD model for collision detection

The first image capture is at a time instance of T_1 and the second capture is at T_{FT}. The captured samples are then compared with the reference voltage and the digital output is stored in the two *SRAM* cells available in the pixel. The differential image of the two spatially integrated digital images obtained from *SRAM1* and *SRAM2* was computed off chip for this version of the sensor.

The temporal decorrelation of the optic flow obtained using the partial charge transfer is shown in figure 7.10. The figure shows the variation in the percentage of active high pixels of two image captures for varying distance of the object from the imager. The first image is captured after an integration time of T_1 and the second after the total frame time T_{FT}. Two measurements for T_1 when 10ms and 1.5s are shown in the figure 7.10. The variation of the time instances allows to generate varying decorrelated one-dimensional binary optical flow.

The difference of the two temporally decorrelated optical flows is plotted in the figure 7.11. It can be observed that as the object moves towards the image sensor, it has a certain threshold of percentage of pixels with changed states below which the object is very near to collision. In this case the collision detection mechanism does not need to use dedicated motion processing blocks. The collision can be detected to a very good degree of reliability using the percentage of changed pixels with the varying one-dimensional differential optical flow.

Harrison [23] uses a collision detection algorithm based on the basic *EMD* model and the algorithm peaks at 230ms before collision for an object with a velocity of 17cm/s. Thus the collision alert is generated at a distance of approximately 4cm.

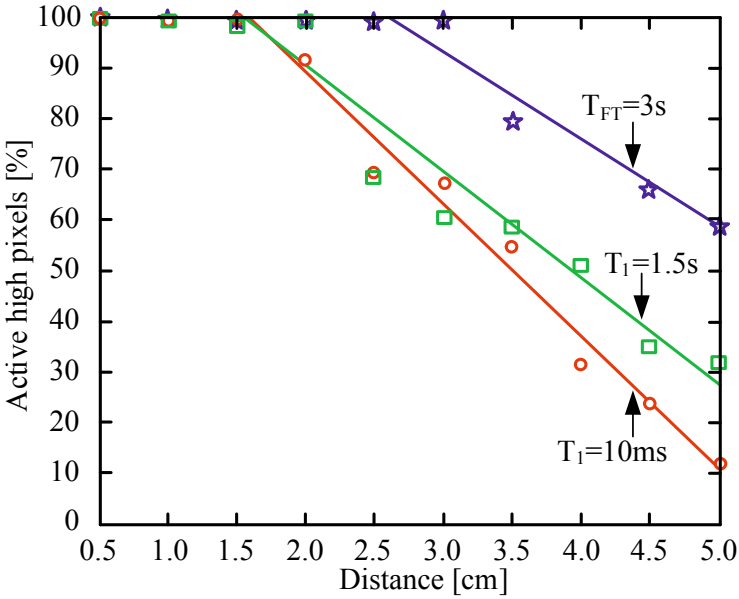


Fig. 7.10 Temporally decorrelated optical flow with approaching objects

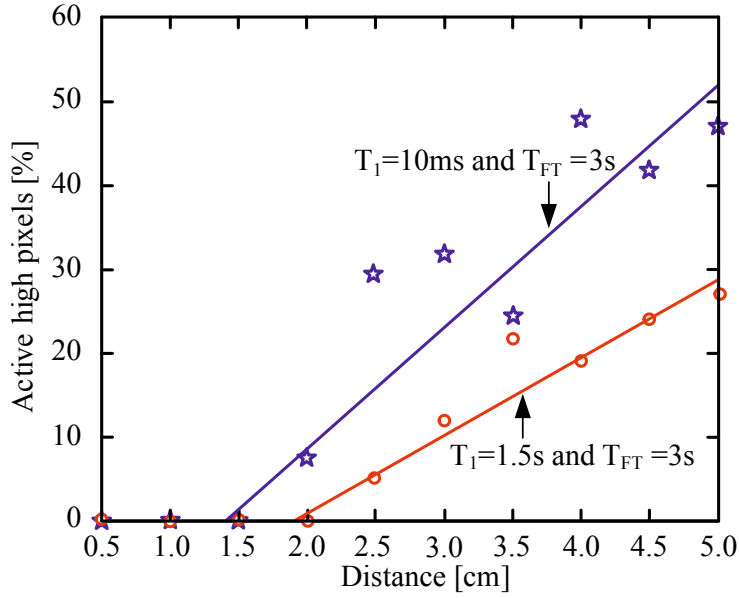


Fig. 7.11 Temporal differential decorrealated optical flow with approaching object

A similar method used by Okuno and Yagi [22] produces a collision detection alert at a distance of 63m. The differential optical flow is shown to generate a collision alert for distances less than 2cm. For narrow path autonomous agent navigations, a collision alert for very small distance is desired. For example in applications like endoscopy, a collision alert distance of less than a cm would be ideal. The distance of collision alert can also be modulated by varying the T_1 and T_{FT}, depending on the application and situation.

Table 7.1 Performance comparison

	[22]	[23]	This work
Method used	*EMD* model based on locust	Delay and correlate *EMD*	Differential optical flow imaging
Collision alert distance	63m	$\sim$4cm	$\leq$2cm

Furthermore the proposed algorithm doesn't suffer from the inherent disadvantage of the *EMD* model in the accurate detection of the output peak. The output of the differential optical flow continues to stay low near collision allowing thresholding and thus is more stable. By modulating the differential time, it would be possible to prevent collision in very narrow paths thus helping navigation of the autonomous agents. Further as most of the computations are done on focal plane there is no image transfer bottleneck, which is usually present in the conventional approach of using an image sensor together with a digital signal processor.

So far the collision detection algorithm is employed in an environment with an illuminated object moving towards the image sensor in a dark background. However since the temporally decorrelated signal is a differential signal, it would also work in an environment condition where a dark object moves vertically with respect to the image sensor in a light background. In such a case initially the binary optic flow would be composed off all digital '1' and as the dark object approaches the image sensor, the optic flow would toggle to digital '0'. The temporally decorelated difference image would thus subtract two images whose optic flow would be dominated by digital '0'. The difference image would thus show the % percentage of active high pixels to exhibit a behavior similar to figure 7.11.

7.3.2 Motion in Horizontal Direction

In this section only horizontal motion is considered. From the pixel array only the binary output of the comparators are used for the motion detection. The 7-bit counter counts the number of '1' in each row of the pixel for each frame. The algorithm for motion detection then compares the counter outputs to decide if there is motion. If the difference of the counter outputs for two exposures is higher than a certain threshold, motion occurrence is flagged.

For the designed sensor the brightness control voltage is the reference voltage to which the analog signal obtained after each exposure is compared. The two *SRAM*

cells in the pixels serve as frame latches and offer both past and current data. The pixel converts the image data into a one-bit data stream by the comparators.

To verify the proposed model two consecutive frames of a light source moving over the image sensor are shown in figure 7.12.

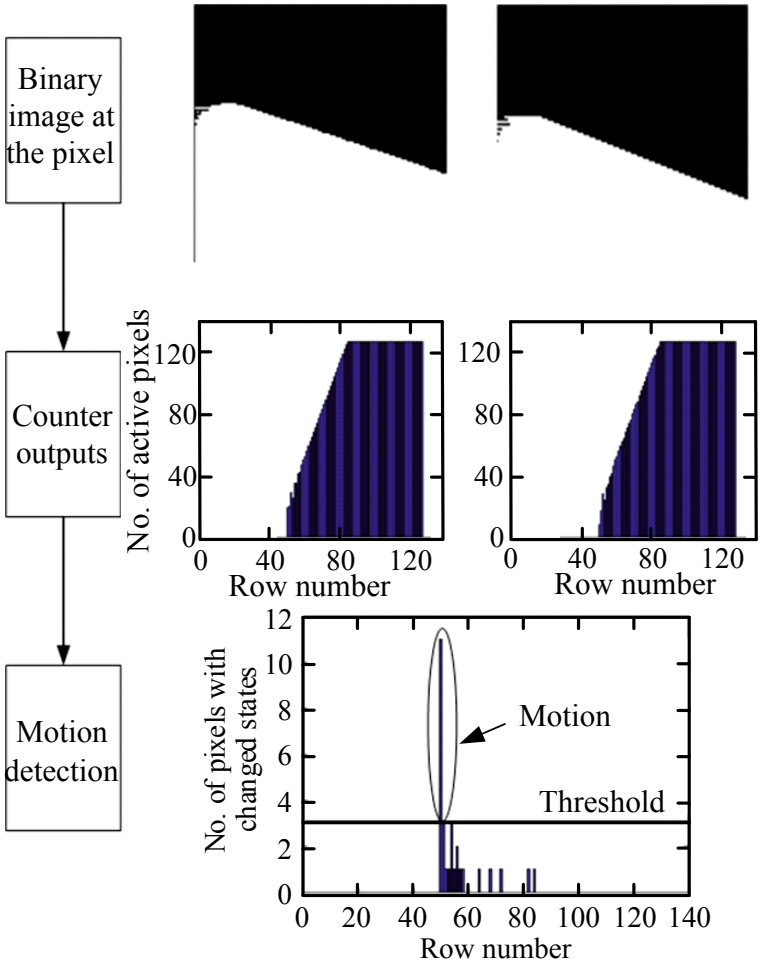


Fig. 7.12 Horizontal motion detection using spatially integrated binary optical flow

The left image shows the light source at its initial position and the right image shows it after a slight movement. The two images look very similar, as only a very small motion was introduced. The histograms of the two images are shown in figure 7.12. The subtraction of the two images results in a difference image, and the histogram of the pixels which changed states are shown at the bottom of the figure. By selecting a proper threshold, accurate detection of motion can be done.

7.4 Illumination Invariant and High Dynamic Range Motion Detection

The outputs of the *EMD* are not invariant to the changes in the background illuminations as discussed in section 7.3. This problem is illustrated in figure 7.13 which shows the increase in the *EMD* output with the increase in the background illumination.

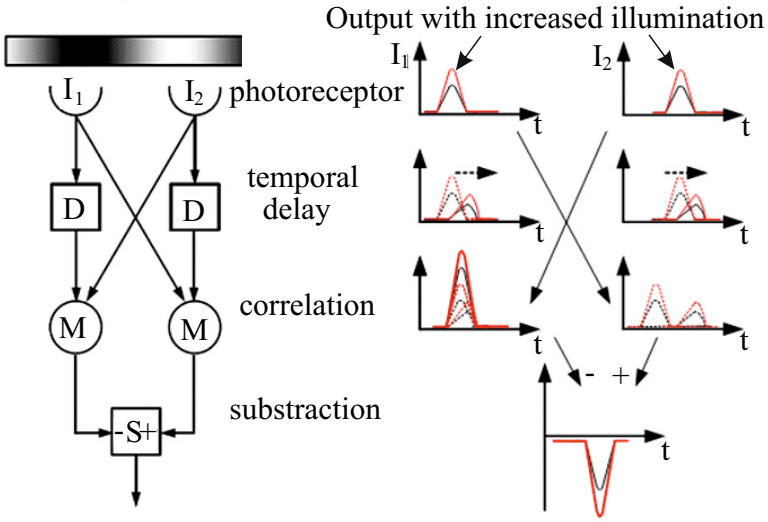


Fig. 7.13 The correlational elementary motion detector showing increased motion for change in illumination

An increase in the background illumination increases the output of the *EMD* and this increased output can be confused with motion. Also it is shown than the optical flow computed on the image doesn't match actual motion with non-uniform illumination [29]. The estimation of motion parameters from the spatiotemporal patterns of visual stimuli or optical flow implicitly requires responses that are invariant with respect to absolute illumination levels, contrast and the spatial structure of the scene.

For the *EMD* model to have a good performance over varying illumination conditions, the dynamic range of the sensor also needs to be high. A real-world scene is composed of varying level of brightness within it. The dynamic range of the scene is typically always higher than the dynamic range of the sensor used to capture the scene. A typical image sensor has a dynamic range of about 65 to 75dB while a scene can have a dynamic range of over 120dB. In high illumination conditions the photodiode of the pixel saturates very quickly. For a very bright object most of the photodiodes will be saturated and thus finer details and motion cannot be captured.

To increase the dynamic range either the maximum signal at the photodiode node (well capacity) has to be increased or the read noise has to be decreased

(equation (3.15)). The relationship between output, integration time and photocurrent at the floating diffusion node of a linear image sensor is given by:

$$V_{signal} = \frac{t_{int} \times I_{photo}}{C_{eff}} \tag{7.14}$$

The photocurrent produced in the photodiode is proportional to the irradiance and the light collection area or the photodiode area. The larger the pixel exposed surface, the more light it can collect during the exposure period and thus the higher is the dynamic range. The photodiode area is limited by pixel size constraints. For high resolution a small pixel size is desired. This leaves us with two parameters in equation (7.14) that can be modified to increase the dynamic range: the integration time t_{int} and the effective capacitance to store the collected charges during integration C_{eff}. The total capacitance C_{eff} on the photodiode sense node is the sum of the junction capacitance proportional to the diode area and the parasitics due to the sensing and amplifying circuitry.

From equation (7.14) it is clear that the slope of the pixel response can be changed by the scaling of the effective capacitance. Changing the effective capacitance varies the well capacity. The effective capacitance can be varied using either smart reset pixels [30], using overflow MOSFET capacitors [31] or by using multiple shorter exposure periods [32], [33]. In the smart reset pixels, the reset gate voltage is monotonically decreased during integration causing the well capacity (charge capacity) of the sensor to monotonically increase. Using a lateral overflow capacitor, the overflowed charges from a fully depleted photodiode during an exposure can still be integrated. The disadvantage of such a scheme is the effective increase of the floating node capacitance which reduces the charge conversion gain and also the fill factor of the pixel. In the multiple shorter integration period methods, several images with different exposure time are captured. The images with a shorter exposure time captures the brightest areas of the scene, while the images with longer exposure time capture the darker areas of the scene. A high dynamic range image is then synthesized from the multiple captures.

Since differential imaging is used in the proposed model of the *EMD* to obtain the temporal decorrelation necessary for motion detection, a partial charge transfer method is used to increase the dynamic range of the model. This method provides a signal after a short integration time in addition to a signal after a long integration time [18]. The difference of the two charge accumulation times in one frame extends the dynamic range of the sensor by removing the static or DC light.

The photo-conversion characteristics for the complete charge transfer *(DDS)* and the differential signal for the partial charge transfer with various integration is shown in figure 7.14. T_1 is varied from 5ms to 20ms.

Figure 7.15 shows the variation in the saturation intensity with varying $T_1 : T_{FT}$ ratio. The saturation intensity is the intensity of light causing the photodiode to saturate. It is observed that the saturation intensity increases with an increase in $T_1 : T_{FT}$ ratio or with an increase in the accumulation time T_1. By appropriately changing the $T_1 : T_{FT}$ ratio, the saturation intensity can be doubled thus increasing the sensitivity to the background brightness.

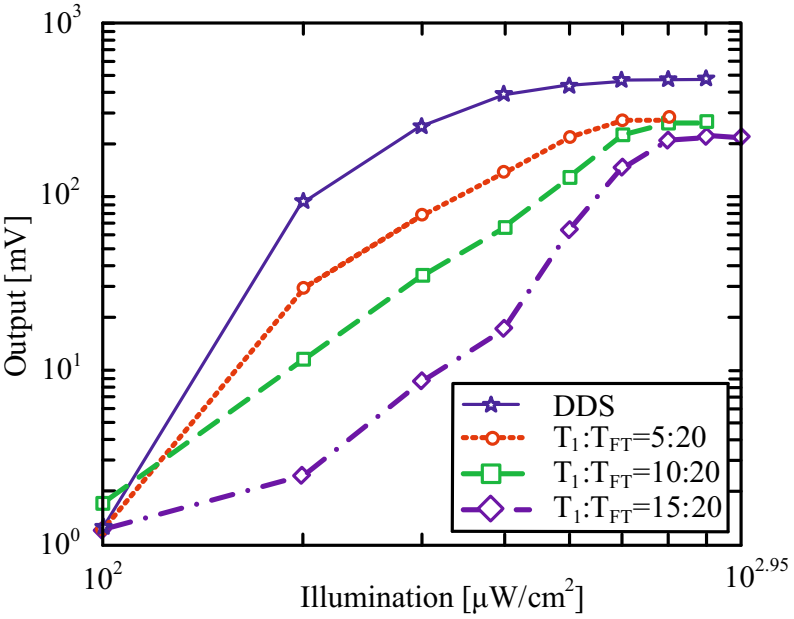


Fig. 7.14 Photo-conversion characteristics at $\lambda = 550nm$

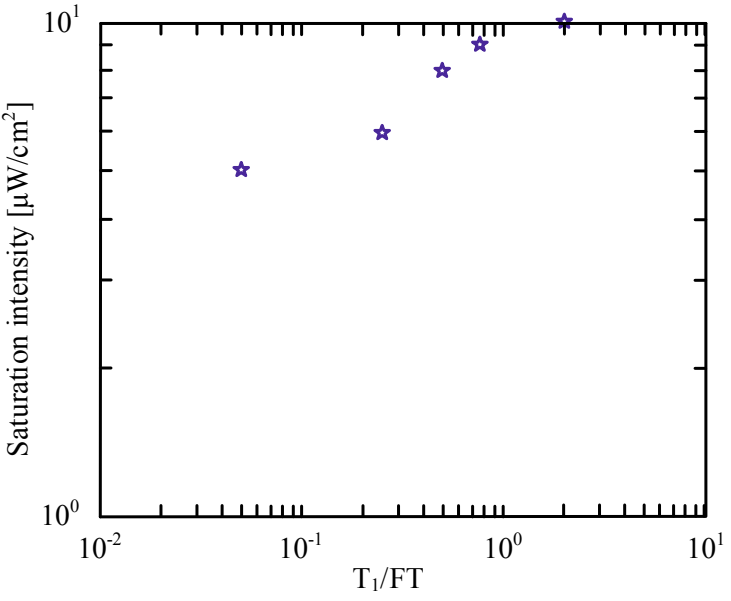


Fig. 7.15 Saturation intensity variation with variation in $T_1 : T_{FT}$ ratio

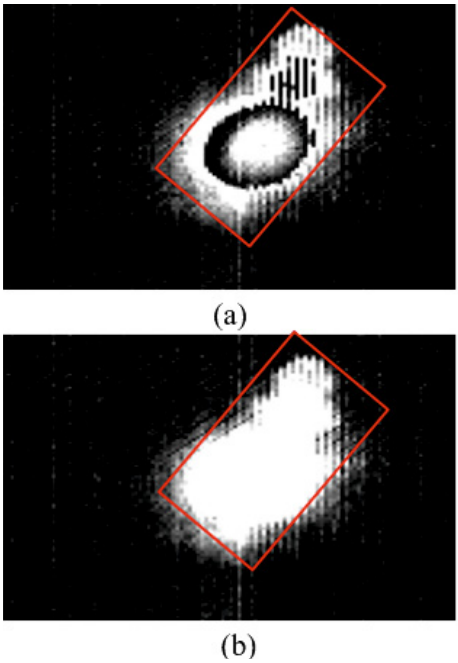


Fig. 7.16 DC light source images with $T_1 : T_{FT}$ of (a) 0.0005 and (b) 0.05

Figure 7.16 shows an example of increasing the saturation level using partial charge transfer. Images of a DC light source were taken with different $T_1 : T_{FT}$ ratios. Figure 7.16(a) and (b) show the images with $T_1 : T_{FT}$ values of 0.0005 and 0.05 respectively. In figure 7.16(a) the pixels at the bright portions of the DC light source saturate, causing an overflow with a very short accumulation time T_1. The differential signal with increased T_1 is shown in figure 7.16(b) and is able to capture the bright portion of the DC light source.

In partial charge transfer, when the charges are transferred from the photodiode to the sense node, some amount of charge is left in the photodiode, as shown in figure 3.24(d) (chapter 3). This introduces image lag and can be removed by resetting the photodiode. However, a reset introduces *kT/C* noise. The amount of residual charge in the photodiode depends on the accumulated charge. Since the accumulated charge is proportional to the exposure time, the residual charge increases with the exposure time. The increase in the residual charge increases in turn the non-linearity of the photo-conversion characteristic and further decreases the sensitivity. The accumulated signal obtained after partial charge transfer with a shorter (T_1) period is also non-linear with respect to the incident light. The non-linearity is believed to be caused by the carrier diffusion and the initial condition of the photodiode [18].

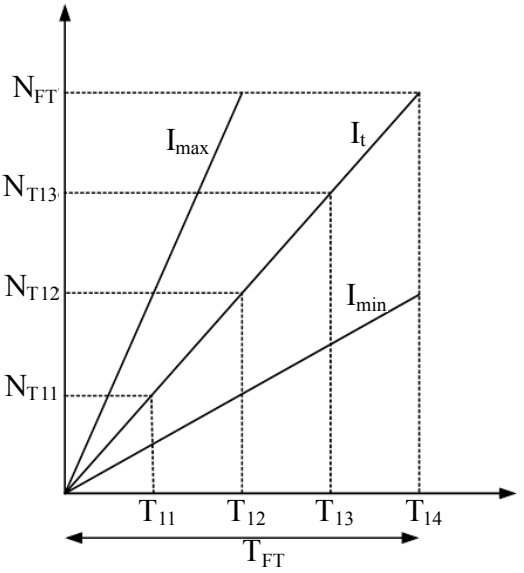


Fig. 7.17 Charge accumulation for various illumination and integration periods

The increase in non-linearity at higher illumination levels for the partial charge transfer can be modeled using figure 7.17. It shows the accumulated charges for various intensities of light I_t (normal), I_{max} (maximum) and I_{min} (minimum) and varying T_1 (T_{11}, T_{12}, T_{13} and T_{14}).

N_{FT} is the number of electrons accumulated when the sensor is operated in conventional or complete charge transfer mode with a frame time of T_{FT}. For the partial charge transfer mode T_1 can vary and four different time instances are shown in figure 7.17. For illumination I_t, the following equations hold

$$q \times N_{T12} = I_t \times T_{12} \tag{7.15}$$

If N_{a1} is the number of residual charges after the first charge transfer at T_{12} for a illumination of I_t then

$$q \times (N_{FT} - N_{a1}) = I_t \times T_{FT} \tag{7.16}$$

$$N_{T12} = \frac{T_{12}}{T_{FT}} \times (N_{FT} - N_{a1}) \tag{7.17}$$

For the highest illumination I_{max}, the photodiode saturates very quickly. N_{a2} is the residual charge after the first charge transfer at T_{12} for an illumination of I_{max}

$$q \times N_{T12} = I_{max} \times T_{11} \tag{7.18}$$

$$q \times (N_{FT} - N_{a2}) = I_{max} \times T_{12} \tag{7.19}$$

$$N_{FT}{}^2 - a_1 \times N_{FT} + b_1 = 0 \tag{7.20}$$

where

$$a_1 = N_{a1} + N_{a2} \tag{7.21}$$

$$b_1 = N_{a1} \times N_{a2} - \frac{T_{FT}}{T_{11}} N_{T12}{}^2 \tag{7.22}$$

Figure 7.18 shows the photo-conversion characteristics of the synthesized wide dynamic range signals with the dynamic range extension ratio T_1 to T_{FT} set to 1:20, 5:20 and 10:20 for higher illuminations. It is observed that for decreasing T_1 to T_{FT} ratios the non-linearity increases at high illumination. It is further observed that with the decrease in the T_1 to T_{FT} ratio the sensitivity increases.

The sensitivities obtained for $T_1 : T_{FT}$ ratio of 10:20, 5:20 and 1:20 are 11.46 mV/ms, 12.89 mV/ms and 14.37 mV/ms respectively.

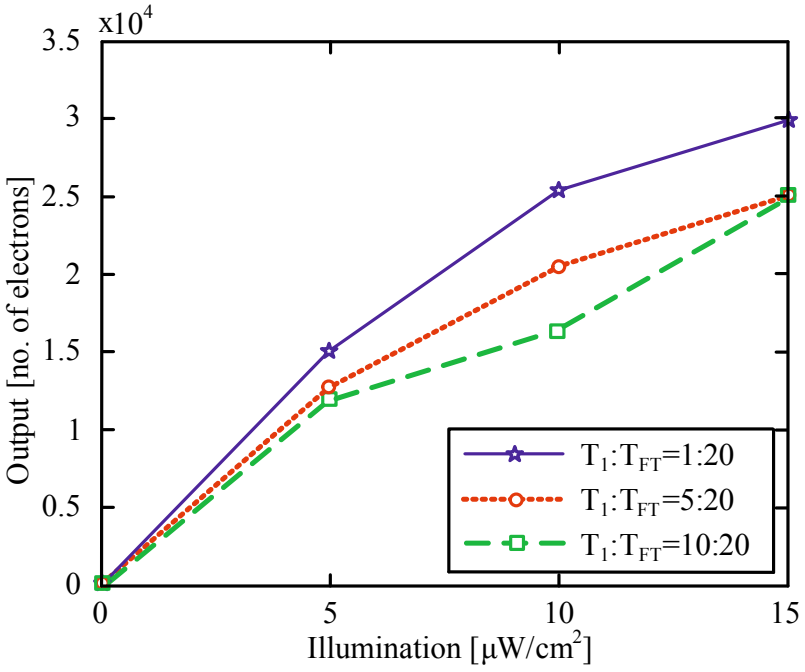


Fig. 7.18 Non-linearity at high illumination for partial charge transfer

The associated non-linearity at high illumination with partial charge transfer is interesting, as in motion detection a non-linear interaction between two spatially separated signals is usually desired [34]. In insects the photoreceptors adapt to the mean luminance in the enviornment and give an approximately logarithmic response to changes in light intensity [35]. This compressive non-linearity allows a better response to contrast than absolute luminance. Non-linearity in the Hassenstein-Reichardt *EMD* model is introduced either by thresholding or rectification to obtain a saturating contrast response curve for motion adaptation [36], [37]. The effect of the non-linearities associated with high illumination obtained with partial charge transfer on motion detection is yet to be explored.

The collision detection algorithm in section 7.3 is designed using one-dimensional binary optical flow. Figure 7.19 shows the percentage of active high pixels when the photodiode is operated in complete charge transfer $(T_1 = 0ms)$ and partial charge transfer $(T_1 = 10ms)$ respectively. It is observed that with partial charge transfer higher illumination levels are captured, thus increasing the digital dynamic range of the image sensor.

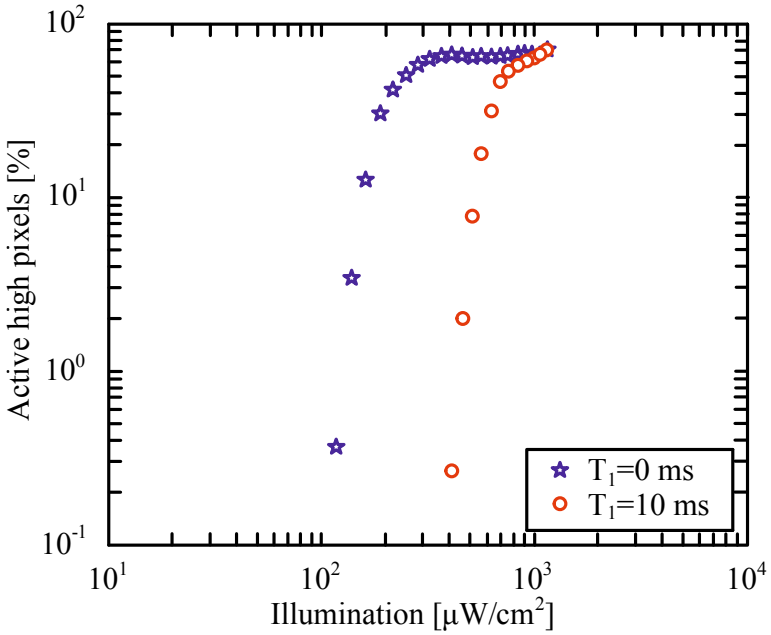


Fig. 7.19 Variations in the optical flow for partial charge transfer

7.5 Digital Polarization

In chapter 4 a CMOS polarization sensor with embedded metallic wire gird was discussed. The image sensor was shown to respond to the changes in the polarization angle. This polarization sensitivity was further shown to be useful in machine vision

applications like material classification and autonomous agent navigation in chapter 5 and 6 respectively. However the proposed solutions for these applications were based on using the analog signals from the pixels. In section 7.3, a simple way to calculate one-dimensional-correlational optical flow using the number of active high pixels at a given time was introduced. The number of active high pixels will vary depending on the horizontal and vertical motion of the object source and its relative intensity variations on the focal plane. Here only motion in vertical direction to the imaging plane is considered. Thus depending on the object moving close or far away from the image sensor, the spot on the focal plane increases or decreases. The size of the spot will be further dependent on the polarization angle of the incoming light ray and the one-dimensional binary optical flow can thus be used to represent polarization in digital format.

The expected theoretical behaviour of the one-dimensional binary optical flow for the two polarization sense regions (chapter 4) is shown in figure 7.20.

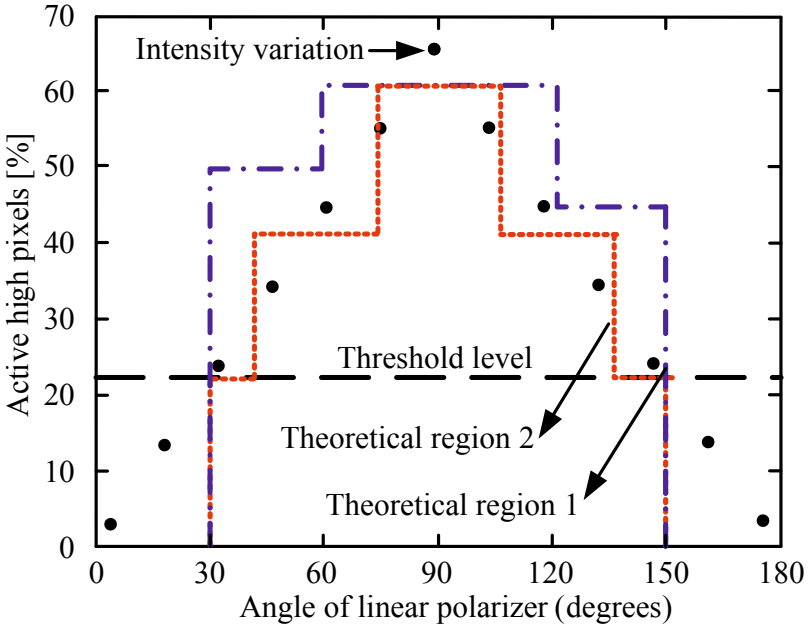
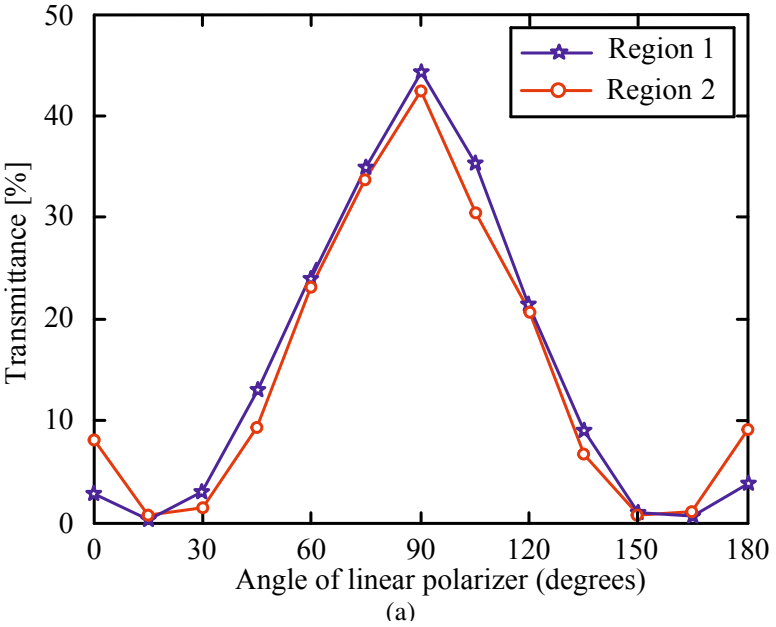
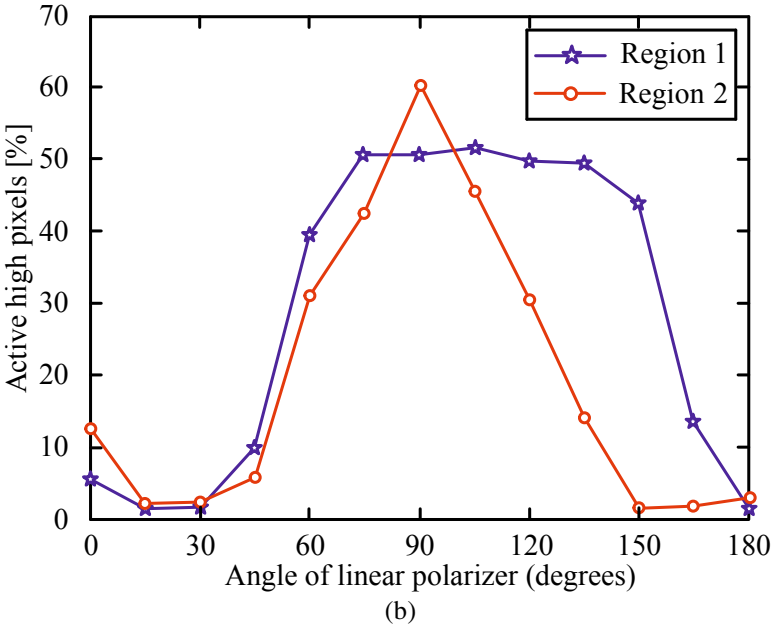


Fig. 7.20 Theoretical behaviour of one-dimensional optical flow with variation in linear polarizer angle

For the polarization sense region 1, when the object is moving towards the imaging plane the intensity of the pixels slowly increases. As the light intensity reaches the threshold level, one expects the two intensity sensitive pixels to store a digital '1', the output of the comparator, in the *SRAM* cells. The percentage of active high pixels shows a step rise to around 50%. As the intensity is further increased around 90°, the 90° sensitive pixels will slowly start to have a have a digital '1',



(a)



(b)

Fig. 7.21 90° polarization in sense regions 1 and 2: (a) analog (b) digital (optical flow variation)

as their comparators' outputs, and hence the percentage of active high pixels will increase above 50%. Ideally, when all the 90° sensitive pixels turn high, 75% of the pixels in region 1 will be active high. However due to the attenuation of the light by the external polarizer not all 90° sensitive pixels turn high. In the polarization sense region 2, the theoretical behavior of the percentage of active high pixels would be the same as region 1, except that as the intensity increases, the 45° sensitive pixels start to turn high. Thus there is an additional step rise when the polarizer angle is 45°. In the experiments, the optical flow is obtained by increasing the light spot gradually from the center to the periphery of the polarization sense regions, using a linear polarizer. This results in a linear increase in the percentage of active high pixels with the variations in the linear polarizer angle, instead of the expected step rise.

The experimental setup for the measurement is the same as described in chapter 4 for analog polarization measurements. The analog performance for the 90° polarizer filter in the two polarization sense regions 1 and 2 is compared with the one-dimensional binary optical flow for varying angle of linear polarizer in figure 7.21.

The measured one-dimensional binary optical flow is shown to have an angular dependence on the angle of the linear polarizer and is very similar to the theoretically predicted behavior. The optical flow and analog representations of polarization in region 2 match closely. It can be predicted that by increasing the number of metallic wire grid orientations over the photodiode a digital representation of polarization very similar to the analog representation can be obtained [38].

A generalized algorithm to represent polarization information will have multiple advantages in low level polarization based machine vision applications. Based on the one-dimensional binary optical flow variations with the polarization angle, a way to determine the Stokes parameters, degree of polarization and polarization Fresnel ratio in binary format can be formulated which will allow focal plane processing of applications like material classification and autonomous agent navigation. Such a sensor would be miniaturized, bandwidth compressor and low power which are highly desirable in the future generations of sensors in machine vision applications.

7.6 Summary

- Motion is usually determined from image sequences. The spatiotemporal image sequences can be represented using the plenoptic function.
- For changing plenoptic function, the intensity of the point object is assumed to be constant.
- Three conventional approaches are used to detect motion: temporal differencing; background subtraction; and optical flow.
- Temporal differencing is based on frame difference. Background subtraction separates the background and foreground and optic flow detects the movement of the brightness patterns on the image plane.
- Optical flow estimation methods can be classified into three main groups: differential methods; matching-based methods and frequency/phase-based.

- The differential methods determines two-dimensional velocities of the moving object employing first or second order spatiotemporal derivatives of the image sequence.
- Matching based methods usually use techniques to either maximize cross-correlation or minimize differential error.
- Frequency/phase based methods use local energy or phase information to determine the velocity of the moving object.
- In motion detection algorithms where the image motion is analyzed by sending the complete frame out of the image sensor, a very high data rate is required to keep the time delay small.
- Temporal differencing at a pixel level rather than at a frame level can help to decrease the data coming from the camera as only the data from the pixel with changed states will be sent out.
- The most popular bio-inspired visual guided motion detector is the correlation-type elementary motion detection (EMD). The optic flow is computed by correlating the response of one photoreceptor to the delayed response of an adjacent photoreceptor, both looking in the same direction.
- The compound eyes of insects are better suited for detection of motion. Inspired by the them, a CMOS image senor operating in temporal differential mode and spatial integration of one-dimensional binary optical flow to detect motion/collision of moving objects was designed.
- The dynamic range of a sensor can be enhanced using partial charge transfer, which also provides for a background illumination invariant motion detection system.
- The one-dimensional binary optical flow also have an angular dependence on the angle of the linear polarizer. By increasing the number of wire grid orientations, a digital representation of the polarization very similar to the analog can be obtained.

References

[1] Adelson, E., Bergen, J.: The plenoptic function and the elements of early vision. In: Landy, M., Movshon, J. (eds.) Computational Models of Visual Processing, pp. 3–20. MIT Press, Cambridge (1991)

[2] Barron, J., Fleet, D., Beauchemin, S.: Systems and experiment performance of optical flow techniques. International Journal of Computer Vision 12(1), 43–77 (1994)

[3] Negahdaripour, S., Horn, B.: Direct passive navigation. IEEE Transactions on Pattern Analysis and Machine Intelligence 9(1), 168–176 (1987)

[4] Anderson, C., Burt, P., van der Wal, G.: Change detection and tracking using pyramid transformation techniques. In: Proceedings of SPIE-Intelligent Robotics and Computer Vision, vol. 579, pp. 72–78 (1985)

[5] Haritaoglu, I., Davis, L., Harwood, D.: W4 who? when? where? what? a real time system for detecting and tracking people. In: International Conference on Automatic Face and Gesture Recognition, pp. 222–227 (1998)

[6] Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7), 780–785 (1997)
[7] Doretto, G., Soatto, S.: Editable dynamic textures. In: IEEE Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 137–142 (2003)
[8] Elgammal, A., Harwood, D., Davis, L.: Non-parametric Model for Background Subtraction. In: Vernon, D. (ed.) ECCV 2000. LNCS, vol. 1843, pp. 751–767. Springer, Heidelberg (2000)
[9] Borst, A.: How do flies land? from behavior to neuronal circuits. Bio-Science 40(4), 292–299 (1990)
[10] Brinkworth, R., O'Carroll, D.: Robust models for optic codeing in natural scences inspired by insect biology. PLoS Computational Biology 5(11), 1–14 (2009)
[11] Tanner, J., Mead, C.: An integrated analog optical motion sensor. In: VLSI Signal Processing II, pp. 59–76. IEEE Press, New York (1986)
[12] Delbrück, T.: Silicon retina with correlation-based, velocity-tuned pixels. IEEE Transactions on Neural Networks 4(3), 529–541 (1993)
[13] Kramer, J.: Compact integrated motion sensor with three-pixel interaction. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(4), 455–460 (1996)
[14] Moini, A., Bouzerdoum, A., Eshraghian, K., Yakovleff, A., Nguyen, X., Blanksby, A., Beare, R., Abbott, D., Bogner, R.: An insect vision based motion detection chip. IEEE Journal of Solid-State Circuits 32(2), 279–283 (1997)
[15] Collins, R.T., Lipton, A.J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., Hasegawa, O., Burt, P., Wixson, L.: A system for video surveillance and monitoring. CMU VSAM Final Report, 1–69 (1999)
[16] Lichtsteiner, P., Posch, C., Delbrück, T.: A A 128×128 120db 30mw asynchronous vision sensor that responds to relative intensity change. In: IEEE ISSCC Digest of Technical Papers, pp. 508–509 (2006)
[17] Innocent, M., Meynants, G.: Differential image senor with high common mode rejection. In: Proceedings of European Solid-State Circuits Conference, pp. 483–486 (2005)
[18] Shafie, S., Kawahito, S., Halin, I., Hasan, W.: Non-linearity in wider dynamic range CMOS image sensors utilizing a partial charge transfer technique. Sensors 9, 9452–9467 (2009)
[19] Srinivasan, M.V., Zhang, S.W., Chahl, J.S., Barth, E., Venkatesh, S.: How honeybees make grazing landings on flat surfaces. Biological Cybernetics 83, 171–183 (2000)
[20] Borst, A.: Correlation versus gradient type motion detectors: the pros and cons. Philosophical Transactions of the Royal Society, B: Biological Sciences 362(1479), 369–374 (2007)
[21] Reichardt, W.: Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. Principles of Sensory Communication, 303–317 (1961)
[22] Okuno, H., Yagi, T.: Bio-inspired real-time robot vision for collision avoidance. Journal of Robotics and Mechatronics 20(1), 68–74 (2008)
[23] Harrison, R.: A biologically inspired analog IC for visual collision detection. IEEE Transactions on Circuits and Systems-I 52(11), 2308–2318 (2005)
[24] Reiser, M., Dickinson, M.: A test bed for insect-inspired robotic control. Philosophical Transactions of the Royal Society, A: Mathematical, Physical and Engineering Sciences 361, 2267–2285 (2003)
[25] Neumann, T.R., Bülthoff, H.H.: Behavior-oriented vision for biomimetic flight control. In: Proceedings of the International Workshop on Biologically Inspired Robotics, pp. 196–203 (2002)

[26] Krapp, H.: Neuronal matched filters for optic flow processing in flying insects. In: Neurological Processing of Optic Flow, pp. 93–120. Academic Press, San Diego (2000)
[27] Tao, S., Zeng, L.: An elementary-motion-detector based on hardware lateral inhibition network. Journal of Physics: Conference Series 48, 212–216 (2006)
[28] Lee, D.: A theory of visual control of braking based on information about time-to-collision. Perception 5, 437–459 (1976)
[29] Horn, B., Schunck, P.: Determining optical flow. Artificial Intelligence 17, 185–203 (1981)
[30] Decker, S., McGrath, R., Brehmer, K., Sodini, C.: A 256 x 256 CMOS imaging array with wide dynamic range pixels and column- parallel digital output. IEEE Journal of Solid-State Circuits 33, 2081–2091 (1998)
[31] Akahane, N., Sugawa, S., Adachi, S., Mori, K., Ishiuchi, T., Mizobuchi, K.: A sensitivity and linearity improvement of a 100db dynamic range CMOS image sensor using a lateral overflow integration capacitor. In: Symposium on VLSI Circuits, Digest of Technical Papers, pp. 62–65 (2005)
[32] Yadid-Pecht, O., Fossum, E.: Wide intrascene dynamic range cmos aps using dual sampling. IEEE Transactions on Electron Devices 44, 1721–1723 (1997)
[33] Yang, D., Gamal, A.E., Fowler, B., Tian, H.: A 640 x 512 cmos image sensor with ultra-wide dynamic range floating-point pixel level ADC. IEEE Journal of Solid-State Circuits 34, 1821–1834 (1999)
[34] Buchner, E.: Elementary movement detectors in an insect visual system. Biological Cybernetics 24, 85–101 (1976)
[35] van Hateren, J., Snippe, H.P.: Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells. Vision Research 41(14), 1851–1865 (2001)
[36] Buchner, E.: Behavioural analysis of spatial vision in insects. In: Photoreception and Vision in Invertebrates. A: Life Sciences, New York, London, vol. 74, pp. 561–621 (1984)
[37] Borst, A., Egelhaaf, M.: Principles of visual motion detection. Trends in Neurosciences 12(8), 297–306 (1989)
[38] Sarkar, M., Segundo, D.S., van Hoof, C., Theuwissen, A.: An analog and digital representation of polarization using CMOS image sensor. In: Proceedings of 5th European Optical Society Tropical Meeting on Advanced Imaging Techniques (2010) ISBN: 9783000305030

Chapter 8
Future Works

In this book the design of a CMOS image sensor inspired by the compound eye of insects, as well as several applications of such image sensor are presented. The sensor is able to detect polarization information using a micro-polarizer oriented at various angles and created using the metal layers in a standard CMOS process. The detected polarization information is shown to be useful in classifying materials as well as detecting incoming light ray directions for use in autonomous agent navigations. Different algorithms used by the eyes of insects to detect motion were also studied and a model for collision detection has been proposed.

In this final book chapter an overview of possible future works are presented based on the previous chapters.

8.1 Future Works

The field of view of an imaging polarimeter is usually limited by the imaging optics. Commercially available photographic and video cameras have a field of view of about $30^\circ - 50^\circ$(horizontal) by 20°-40° (vertical) depending on the focal length and aperture. Ideally 180° field of view imaging polarimetry is desired to study the polarization patterns.

A 124° field of view camera with less than 1° angular resolution has been discussed in chapter 2 and 3. Such a camera would help to increase the sensitivity of the polarization measurements while allowing to gather wide field of view information.

8.1.1 Wide Field of View Imaging System with Polarization Sensitivity

Wide field of view imaging systems can potentially improve applications such as security surveillance, robot navigation or endoscopy. Such a system can collect image data points from a relatively larger area than the rectilinear lenses and thus can be employed in both contact and contactless endoscopy. The fish eye lenses which

M. Sarkar and A. Theuwissen: *A Biologically Inspired CMOS Image Sensor*, SCI 461, pp. 247–252.
DOI: 10.1007/ 978-3-642-34901-0_8 © Springer-Verlag Berlin Heidelberg 2013

are generally used in endoscopy for images produce a significant distortion in the visual content. Contactless endoscopy is becoming the standard in the endoscopy world. The image sensor with wide field of view, high angular sensitivity and polarization sensitivity would enhance the application areas of the endoscopes.

The common requirements for a wireless capsule endoscope are high resolution, very low power, high dynamic range and high sensitivity. Resolution is needed to be able to see things clearly inside the digestive tract or the intestines. The state-of-the-art commercial wireless capsule endoscope product, the PillCam capsule, developed by Given Imaging Ltd transmits 256x256x8 bit data at a frame rate of 2 frames per second [1]. The sensor described in chapter 3 is able to transmit 128x128x12 bits of data at a frame rate of 15fps. Besides resolution, endoscopy applications demand a dynamic range of 60-80dB which is easily achievable. The dynamic range can be varied using the partial charge transfer as seen in chapter 7. The ability to do in-pixel processing of the data helps in data compression, which naturally helps to lower the output data rate and lower the power consumption of the system.

The ability of the designed sensor to detect polarization of the reflected light adds value to medical imaging. It is known that when light interacts with tissue, the reflectance, absorption and polarization of light is affected. The polarization parameters of light scattered from biological tissues contains rich morphological and functional information useful for medical purposes. Potential applications of such electro-optical system include imaging of superficial cancers and other skin lesions, early detection of diseased cells and microscopic analysis of tissues [2].

A potential application of linear polarization imaging techniques in dermatology has been proposed by [3]. This technique records a series of images corresponding to different combinations of illumination and polarization and calculates the intensity difference between orthogonal direction polarizations pixel by pixel. The degree of polarization information obtained from the surrounding muscles or blood can also be evaluated. In chapter 4, it was shown that it is possible to detect polarization information using metallic wire grids and it is also possible to compute the orthogonal polarization intensity difference and degree of polarization of polarization pixel by pixel. For the current version, these calculations are done off-chip but they can be easily incorporated onto the chip, which will further help in miniaturization of the endoscopes.

Besides endoscopy, wide field of view optical systems can also be used in imaging fluorescence microscopy. Fluorescence microcopy is used to investigate dynamic phenomenon in cells and living tissues and have seen dramatic increase in life science application in the last decade. Conventional CCD/CMOS cameras are not able to image the time-critical cellular events which change dynamically in less than a few seconds. To be able to capture these fast changes, multiple channel imaging optics are required to simultaneously track changes in the fluorescence signals. Besides fluorescence microscopy, fluorescence polarization imaging is also becoming popular in investigating the biochemical properties of samples, such as protein denaturation, protein-ligand interactions, protein-protein interactions, protein-DNA interactions, and the rotational rates of proteins [4], [5]. A wide field of view optical system with polarization sensitivity is ideal for biological

investigations because it provides the ability to image simultaneously both the parallel and perpendicular polarization components of a fluorescence emission.

8.1.2 High Angular Resolution Imaging System with Polarization Sensitivity

In chapter 4, it was elaborated that the performance parameters of a wire grid micro-polarizer depend on the pitch used. To obtain a high transmission efficiency and extinction ratio *(ER)* in the visible region, the pitch has to be smaller than the incident wavelength. The 0.18μm CMOS technology used to fabricate the designed image sensor provides a minimum metal spacing of 0.24μm. Nevertheless, as CMOS technologies keep scaling and smaller feature sizes become available, it is possible to fabricate the suggested structures with smaller spacings.

The polarization sensitivity also depends on the wavelength of the incident electromagnetic wave and the angle of incidence. As wavelength and grating spacing becomes nearly equal, the angle of incidence of the light ray has more pronounced effects on the polarization detection sensitivity. In the case of a wide field of view lens with 1°angular resolution, it offers two advantages. Firstly, the polarization ratio for a specific angle of incidence is known and secondly, it is possible to spatially sum the polarization response over a wide field of view. For the current version of the sensor, the polarization profile was measured for only a single angle of incidence; however after the integration of the lens the polarization measurement for the entire field of view can be obtained which will increase the resolution and sensitivity of the polarization measurement.

8.1.2.1 Real Time Material Classification

In chapter 5, it was shown that it is possible to classify materials based on the polarization information from the reflected surface. To keep the system simple, the degree of polarization and *PFR* calculations were done off-chip. These calculations can be easily implemented on-chip, allowing for real time material classification. Further the transmitted intensities of the reflected light were measured for a single angle of incidence of the light ray. The measurement of the polarization Fresnel ratio *(PFR)* at each pixel is not very accurate when the diffuse component of the reflection dominates over the specular component. In such an scenario the measurement of the reflected transmitted intensities for varying angle of incidence of the light ray would serve to increase the resolution and the sensitivity of the *PFR* measurement.

The multichannel imaging system with an angular resolution of 1°would provide the polarization information of the reflected light over a wide field of view, thus serving as a more reliable measurement and classification system. The increased angular resolution with corresponding increase in the number of Fresnel reflection coefficients can provide a more intuitive way of investigating the specular and diffuse reflection changes as a function of physical parameters of the reflecting surface.

Furthermore, in this work only the differential changes in the intensities of the parallel and perpendicular component of the specular and diffuse reflections were explored and not the phase difference. Metals alter the phase of polarization of the incident light upon specular reflection while dielectrics do not [6]. This is important in the scenario of highly diffuse dielectric materials, which would often be misclassified as metals when using *PFR*. Ways to determine the change in phase of the incident and the reflected light by the reflecting surface need to be explored.

8.1.2.2 Real Time Navigation

In chapter 6, it was stated that insects use the egocentric form of navigation which computes the home vector from the direction of travel and the distance travelled. The possibility of using polarized light as a compass to determine the direction of travel was shown. The distance of travel computation were not discussed, however they are not very difficult to compute. One such method is proposed here which can be easily implemented on chip for distance of travel calculations.

Assuming the initial position of the object to be *(x,y,t)* then if the point has travelled to a new location after a time Δt the new location X and Y can be described as

$$X(t+\Delta t) = x(t) + \cos\theta(t) \times v(t) \times \Delta t \tag{8.1}$$

$$Y(t+\Delta t) = y(t) + \sin\theta(t) \times v(t) \times \Delta t \tag{8.2}$$

where θ is the angle of travel and *v(t)* is the velocity of travel. Δt is the integration time to obtain the next frame information. The velocity can be computed from the optic flow as discussed in chapter 7. However the computation of the velocity from the optic flow would be computationally demanding. The velocity of travel is proportional to the change in the intensity of the image over time [7] and is expressed as

$$\Delta i \approx \Delta v \times \Delta t \tag{8.3}$$

where Δi is the change in intensity over the time interval Δt. Substituting equation (8.3) in equations (8.1) and (8.2) we get

$$X(t+\Delta t) = x(t) + \cos\theta(t) \times \Delta i \tag{8.4}$$

$$Y(t+\Delta t) = y(t) + \sin\theta(t) \times \Delta i \tag{8.5}$$

The new position is depended on the angular travel and the change in intensity during the time Δt. The change in intensity can be computed from the difference in two frames separated by Δt using partial charge transfer as discussed in chapter 7, while the angular information can be obtained from the compass discussed in

chapter 6. Thus it would indeed be possible to have both the direction and distance information which would simplify the algorithms currently used for autonomous agent navigation.

8.1.2.3 Real Time Sun Position Detection

The proposed design will make it possible for an alternate method of angular position determination of the sun besides the slit sun digital sensors. The celestial compass based on skylight polarization is described in chapter 6. Depending on the angle the light from the sun hits the atmosphere; the light is polarized in a special way. These patterns maintain an interesting characteristic all over the day: they are always perpendicular to the solar meridian. Thus the degree of polarization has a maximum when the sun is at 90°angular position (with respect to solar meridian it is at 0°). It was shown that it is indeed possible to determine the elliptical and azimuthal position of an object based on the polarization patterns of the incoming light ray.

Due to the measurement setup constraints, it was not possible for the first version of the designed image sensor to test for the variability of the degree of polarization with respect to the angular position of the sun under open sky. The measurement setup can be simplified and an open sky measurement should be a valuable extension of the project. The study of the change of the degree of polarization with the position of the sun would immensely help in the navigation of autonomous agents which usually depend on heavy computations to determine the angle of travel. Further it is known that sun moves at a constant angular velocity. The change from the solar

Table 8.1 Comparison of sun position detection sensors

	GPS	Magnetometer	SSC [8]	proposed
Mass	∼20gm	∼50gm	∼20gm including processor	same as SSC
Power	∼0.5W	∼0.3-0.5W	∼0.2W	∼0.1W
Acquisition time	1 min	1 sec	10 min	10ms-1sec
Accuracy	0.15°-0.5°	0.5°if position is known	0.1°	Minimum of 1°
Operation on moving vehicles	Yes	Yes	No	Yes
Operational constraint	GPS coverage	Initial position must be known. Will not operate in polar regions	Needs to see the sun	High extinction ratio for increased polarization sensitivity

meridian (the point of highest degree of polarization) and the calculated orientation would also be able to predict the time elapsed.

A comparison between different sensors to determine the sun position is shown in the table 8.1 [8].

References

[1] Metzger, Y., Adler, S.N., Shitrit, A.B., Koslowsky, B., Bjarnason, I.: Comparison of a new pillcam sb2 video capsule versus the standard pillcam sb for detection of small bowel diseases. Reports in Medical Imaging 2, 7–11 (2009)

[2] Ghosh, N., Soni, J., Wood, M., Wallenberg, M., Vitkin, I.: Mueller matrix polarimetry for the characterization of complex random medium like biological tissues. Pramana - Journal of Physics 75, 1071–1086 (2010)

[3] Nan, Z., Xiaoyu, J., Qiang, G., Yonghong, H., Hui, M.: Linear polarization difference imaging and its potential applications. Applied Optics 48(35), 6734–6739 (2009)

[4] Elgass, K., Caesar, K., Schleifenbaum, F., Stierhof, Y., Meixer, A., Harter, K.: Novel application of fluorescence lifetime and fluorescence microscopy enables quantitative access to subcellular dynamics in plant cells. PLoS ONE 4(5), e5716-1–e5716-13 (2009)

[5] Mashanov, G., Nenasheva, T., Peckham, M., Molloy, J.: Cell biochemistry studied by single molecule imaging. Biochemical Society Transactions 34(5), 983–988 (2006)

[6] Chen, H., Wolff, L.: Polarization phase-based method for material classification in computer vision. International Journal of Computer Vision 28(1), 73–83 (1998)

[7] Lopez, L.: Neural processing for control of artificial compound eye. In: IEEE International Conference on Computational Intelligence, vol. 5, pp. 2749–2753 (1994)

[8] Liebe, C.C.: Solar compass chip. IEEE Sensors Journal 4(6), 779–786 (2004)

Summary

Biological systems are a source of inspiration in the development of small autonomous sensor nodes. The two major types of optical vision systems found in nature are the single aperture human eye and the compound eye of insects. The latter are among the most compact and smallest vision sensors. The visual system of insects allows them to fly with a limited intelligence and brain processing power. A CMOS image sensor replicating the perception of vision by insects is discussed and designed in this book for industrial (machine vision) and medical applications.

Visual perception is the ability to detect light and interpret it. In chapter 1 of the book different vision perceptions found in the animal kingdom are briefly described. Among the known optical vision systems in nature, the single aperture eye is more popular with camera researchers for its ability to form high resolution images. Compound eyes are typically not suitable for high resolution images. However they are known to process information more efficiently and are good examples of low power vision systems. Light emanating from reflections is partially polarized. Single aperture eyes are polarization blind, but compound eyes are able to detect polarization. The compound eyes are also more suited for motion detection, have high sensitivity in low light conditions and offer wide field of view at a cost of low optical resolution. Though the compound eyes are not the best in terms of image quality, the advantages that they offer makes them better suited for machine vision applications than the single aperture eyes.

Compound eyes are composed of many ommatidia or facets. The natural compound eyes are classified broadly into apposition and superposition eyes. In apposition eyes each facet receives light only from one direction while in superposition eyes the light rays from different facets are superimposed. To design an artificial compound eye inspired by the natural compound eye, a micro-optical lens array and a detector array are required. In chapter 2 and 3, a realization of a single camera system wherein the scene is optically distributed across multiples lenses using a single detector array is presented. To optically distribute the scene, a compound lens structure has been designed by Vrije Universiteit Brussel *(VUB)*

within the project of which this work is also part of. The designed lens has a field of view of 124°, a maximum contrast of 0.3 line pairs per degree (LP/°) and an angular resolution of less than 1°.

In chapter 3, the design of the detector array for said optical system is presented. The designed image sensor consists of an array of 128 by 128 pixels, it occupies an area of 5mm x 4mm and it has been designed and fabricated in the 0.18μm CMOS CIS process from UMC. The pixel used in the design is a smart pixel, which is a form of an active pixel sensor. Each pixel contains a pinned photodiode $(p^+/n/p)$ and 33 transistors to perform low level image processing. The size of the pixel is 25μm x 25μm and the size of the photodiode is 10μm x 10μm which corresponds to 16% fill factor. The pixel also has an in-pixel analog and digital memory, which enables in-pixel digitization and real-time dynamic range adaptation. The designed image sensor can be operated in two modes: double differential mode and differential imaging mode. In case of double differential mode each pixel samples the voltages at the photodiode node twice: once after reset and again after the exposure period. In the differential mode two samples are obtained after different exposure periods.

In the designed image sensor only half of the available pixel array of 128 by 128 is used for normal imaging. The other half of the pixel array is used to analyze polarization information using metallic wire grid structures discussed in chapter 4. The three basic characteristics of light are intensity, color and polarization. Polarization vision and the mechanisms of the polarization of light from unpolarized light beam along with the representation of the polarization information are also discussed in detail in chapter 4. Polarization vision is a generalization of intensity and is analogous to color vision. In section 4.5 a metallic wire grid micropolarizer operating in spatial mode is presented which is designed using the metal layers available with the standard CMOS technology. Normally such a wire grid structure would function as a simple diffraction grating but when the pitch of the wires is less than half the wavelength of the incoming light, it becomes a polarizer. The designed polarization sensing sensor has an embedded linear polarizer in each pixel with varying orientations. The linear wire grid polarizer is implemented using thin metal strips with a line/space of 240nm/240nm (pitch of 480nm). The absorption of the EM waves to completely *s*-polarize or *p*-polarize the transmitted wave through the wire grid is dependent on the pitch of the wire grid and so the performance of the wire grid as a polarizer also varies with the pitch. With the scaling of the CMOS technologies, the pitch of the wire grid will also scale thus improving their performance as a polarizer.

The pixel array used for polarization sensing is further split into two 64 by 64 pixel regions designated sense regions 1 and 2. In sense region 1, the wire grid is oriented horizontally and vertically to measure the 0° and 90° polarization intensity. While in sense region 2 beside the horizontal and vertical orientations the wire grid oriented in 45° is also present. Linear polarizers are characterized by two main specifications: transmittance and extinction ratio. In the sense region 1, for 0° polarization the maximum and the minimum transmittances observed are 38.9% and 7% respectively. For 90° polarization the maximum and the minimum

transmittances observed are 44.4% and 0.1% respectively. Similarly, in sense region 2 for 0° polarization the maximum and the minimum transmittances observed are 38.4% and 5.4% respectively while for 90° polarization the maximum and the minimum transmittances observed are 42.4% and 0.6% respectively. The extinction ratios obtained in the sense region 1 and 2 are 6.3 and 7.7 respectively. These extinction ratios are higher than reported in literature for configuration similar to the designed sensor.

In chapter 5, material classification using polarization information is presented. Materials are broadly classified into metals and dielectrics. The classification of the material type can provide vital information about the scene in computer or machine vision. Unpolarized light becomes partially polarized after specular reflection. The state of polarization for the diffuse and specular components of the reflection depends on the reflecting surface and the measurement of the state of polarization of the reflected light serves as an indicator for the type of material surface. The magnitude of oscillations of the maximum and minimum transmitted irradiance due to the variations in the reflection pattern of the metal and dielectric surface was shown to be useful in classifying them. Various other measurement metrics, such as the degree of polarization and the polarization Fresnel ratio were also shown to measure the variations in the reflection pattern of the metal and dielectric surface. The Stokes degree of polarization was also evaluated to classify materials. The polarization of the reflected light also varies with the conductivity of the metallic surface and this was further explored and was shown to be able to serve as a tool to classify among highly conductive and lowly conductive metallic surfaces.

In chapter 6, navigation using polarized light is discussed. The navigational strategies of insects using skylight polarization are interesting for applications in autonomous agent navigation because they rely on very little information. The polarization pattern in the skylight varies in a systematic fashion both in *e*-vector and degree of polarization, according to the position of the sun. The Stokes parameters are used to calculate the ellipticity and azimuthal angles of a Poincaré sphere and these angles are shown to be correlated to the incoming light ray direction. A correlation coefficient higher than 0.94 was obtained in all the measurements. This method thus provides a way to determine the incoming polarized light direction which could be used to determine the angular positional information useful for autonomous agent navigation. The ability to compute on-chip or in real-time the positional information would result in highly miniaturized navigational sensors and saving computational power. Further the degree of polarization was shown to vary with the polarization of the incoming light ray direction which could also serve as a compass for autonomous agent navigation algorithms. The variation in the degree of polarization with the position of the sun can further be explored in sun position detection based on skylight polarization.

In chapter 7 the ability of the insect to detect fast motion is explored and also the concept of representing polarization information in binary format is presented. The ability of flying insects to detect fast motion in the visual scene and avoid collision using low level image processing and little computational power makes their visual processing interesting for real time motion/collision detection in machine

vision applications. The designed image sensor can also be operated in temporal differential mode and spatial integration of one-dimensional binary optical flow to detect motion/collision of moving objects replicating the "flickering effects" of the insects. The binary optical flow is generated in-pixel from multiple images stored in the in-pixel memories and spatially integrated using a counter. Both horizontal and vertical motions are easily detected using the binary optical flow which is computed relatively easily in the pixel. A collision alert algorithm to an approaching object as close as 2 cm is demonstrated. Such a collision detection algorithm would allow for the design of simple, miniaturized, low power and narrow path autonomous navigating agents. The designed sensor when operated in differential imaging mode provides for a background illumination invariant motion detection system. The generated binary optical flow is shown to have an angular dependence to the linear polarizer and with increasing the wire grid orientations the digital representation of the polarization resembles the analog form.

In chapter 8, along with the summary of achievements a multichannel imaging system was discussed. The integration of the compound lens from *VUB* with the designed detector array would result in a multichannel imaging system. With high field of view and a low angular resolution, these optical vision systems would find applications not only in machine vision but also in medical applications.

Acknowledgments

"You can complain because roses have thorns, or you can rejoice because thorns have roses", exactly summarizes how you want to look into the long years spent on this book. This book is an extension of my PhD thesis. It is so very obvious to be influenced by many during the course of these seemingly never ending years.

There are many whom I want to thank for helping me finish with my thesis. I will start with my promoter Albert Theuwissen, who has been very supportive and encouraging over these years. I am really thankful to him for accepting me as a PhD student even with me not having any background knowledge on image sensors and for being a constant source of motivation.

I would also like to thank Imec, Eindhoven for financing the project and particularly Bert Gyselinckx who offered me the PhD position at Imec and Margot Nijkamp who took care of the official requirements. For the New Year celebrations, I was gifted twice by Holst, the normal Christmas gift and a new supervisor to work with which made life more interesting. I started the project with Guy Meynants and I am immensely grateful to him for imparting to me his knowledge on image sensors. I designed my first working chip under his supervision and it will remain an unforgettable achievement for years to come. Patrick Merken taught me to be patient and it was fun to have him around. If my technical writing ability has improved the credit has to go to San Segundo Bello David, for his constant guidance on this aspect. He has been very supportive of my work and has always motivated me. David has also immensely helped me in writing my thesis.

I would also like to thank Chris Van Hoof for all the travel budgets for the conferences that I attended. I imagined Chris declining the idea every time I approached him for a conference attendance but he never did. The exposure that I got from the conferences was really very helpful for my scientific and personal growth, and I am really indebted to him for that. Thank you, Chris.

The office colleagues in Imec, Eindhoven have always supported me in my work. I would particularly like to mention a few names, Jiawei with whom I had lots of discussions on amplifiers and general analog circuit design. Besides circuit design we often ended up talking about the share markets, one passion that we both shared and it helped that we were in the same office. Richard van der Hoven, my coffee

partner and Friday evening party mate. I always had a company to go out with and that helped to charge up after a long week at work. Thanks to Richard and Pieter Harpe for helping me to translate my summary and propositions to Dutch. All the other office colleagues were also supportive always. Amongst all, I would like to particularly acknowledge Phillips Mattelaer, the Business Development Manager of Holst. Occasionally he would drop by my office and would ask his famous question "Can your facet eye do this?" to which I always replied "Sure, why not". Thank you Phillip for all the discussions we have had on crazy technologies possible using the insect's vision and for always motivating me to think for ideas and possibilities.

I would also like to sincerely thank Adri Mierop of Dalsa for allowing me to use his experimental setup for measurements and helping me with the debugging process that eventually brought the sensor to life. The thesis writing has been a long journey and thanks Mini for helping me with it.

I would also like to thank Hans Boeve and Martin Ouwerkerk from Philips research, Eindhoven who helped me to move to Netherlands from Germany for my master thesis and later was when applying for a PhD position at Holst. Particularly I want to thank Hans who showed a lot of confidence in me and encouraged me to go for a PhD in the first instance.

And last but not the least, I would like to thank my family who has been a pillar of support all these years. I particularly thank my Mom who supported me in all my engagements and had confidence in me all the way even when others doubted. Love you, Mom.

New Delhi, October 2012 Mukul Sarkar